MED. CAL. LIBRARY

HUMAN BIOLOGY

a record of research

DECEMBER, 1947

Vol. 19 No. 4

PUBLISHED QUARTERLY BY
THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U.S.A.

HUMAN BIOLOGY

a record of research

FOUNDED BY RAYMOND PEARL

Edited by

CHARLES P. WINSOR

in association with

A. W. FREEMAN A. D. LANGMUIR ADOLF MEYER
CURT P. RICHTER ADOLPH H. SCHULTZ

Managing Editor
FRANCES D. NICOLLS

ADVISORY BOARD

ED. CLAPARÈDE
University of Geneva

S. R. DETWILER
Columbia University

EUGEN FISCHER
Kaiser Wilhelm Institut für Anthropologie, Berlin-Dahlem

CORRADO GINI
R. Università de Roma

MAJOR GREENWOOD
School of Hygiene and Tropical Medicine, London

SHINKISHI HATAI
Tôhoku Imperial University, Sendai

ERNST KRETSCHMER
Universitäts-Nervenklinik, Marburg

A. J. LOTKA
Metropolitan Life Insurance Company

D. PEYRONY
Musée Préhistorique, Les Eyzies (Dordogne)

P. RIVET
Universidad de Bogotá, Colombia

JOHN B. WATSON
New York

CLARK WISSLER
American Museum of Natural History

HUMAN BIOLOGY is published quarterly by The Johns Hopkins Press, Baltimore, Maryland, U. S. A. The advance subscription price is $5.00 a year (foreign postage 35 cents; Canadian 13 cents). Single current issues are $2.00 each; back numbers, $3.00 each; back volumes, $6.00; plus postage. Subscriptions for less than a year are charged at the single issue price.

Subscribers should notify the publishers of change in address at least two weeks before publication of the issue with which change is to take effect. No claim for non-receipt of an issue will be entertained unless made immediately upon receipt of the next succeeding issue.

Manuscripts and communications regarding editorial matters should be addressed to Editor, Human Biology, 615 N. Wolfe St., Baltimore—5, Maryland, U. S. A.

THE JOHNS HOPKINS PRESS, *Publishers*

BALTIMORE—18, MARYLAND, U. S. A.

Entered as Second Class matter, February 18, 1929, at the Post Office at Baltimore, Md., under the Act of March 3, 1879.

Copyright 1948 by The Johns Hopkins Press

HUMAN BIOLOGY
a record of research

DECEMBER, 1947

VOL. 19 No. 4

QUANTITATIVE ASPECTS OF ORGANIZATION

BY C. C. LIENAU

TABLE OF CONTENTS

1. Statement of the problem
2. Pioneering attempts to define organization in quantitative terms
3. Groups and some of their order relations
4. Multiple group as a dual group
5. Multiple group
6. Group algebra ($m = 1$)
7. Extension of the algebra to multiple groups ($m = 1, 2, \ldots$)
8. A role for probability theory in contest
9. Integration: its degrees and levels
10. Differentiation of the dual group
11. Integration and differentiation of the multiple group
12. Organization measure
13. Organization and time
14. Representation of unanimity and dissent in voting groups

1. STATEMENT OF THE PROBLEM

CRITICAL study of certain scientific, technological and speculative writing suggests that the idea of organization is overdue to be more clearly formulated. Such clarification must extend to formulation of its possible metrical aspects.

Organization is usually considered to be a fundamental biological or social concept, to have metrical aspects perhaps, but to be mainly relational or qualitative in content. However, it has been suggested

recently that it may be an overworked philosophical idea into which a wide range of scientific and other content can be fitted.[1] Indeed, it has played no fundamental part in the mathematical-physical biology developed since 1900 by Lotka [2] and is only incidental to the biophysics of Rashevsky.[3] The rather too free extension of its usage [4, 5] in the sciences (physical, biological and social) and even in technology (e. g. Kron) as well as in the speculative literatures, can be traced proximately to the 19th century, mainly to Spencer and the organic analogy.[6] Ralph

---

[1] Waddington, C. H., Organisers and Genes, *Cambridge*, 1940, p. 142. ". . . in recent years, there has been a tendency either to regard organisation as one of the irreducible fundamental bases of all biology, or to invoke it, as though it were a well-defined concept, to fill up any awkward gaps in a theoretical structure. The latter use . . . is obviously dangerous without a clear idea of exactly what is meant by the term, and when an attempt is made to reach such clarity, I think it will be found that the scope of the notion, as an explanatory principle, is not as great as has sometimes been suggested. . . . , it requires more precise formulation than it has received, at least in scientific contexts. . . . it is clear that the fundamental notion is organisation, which is capable of quantitative variation, rather than organism, which is not . . . the concept of organisation is usually invoked . . . to form part of a theoretical system for dealing with phenomena which seem to involve the subjection of otherwise self-sufficient parts to some over-riding whole. . . . Where does the organisation come from? . . . What is meant by a higher level of organisation? . . . a new level of organisation cannot be accounted for in terms of the properties of its elementary units as they behave in isolation, but is accounted for if we add to them certain other properties which the units can only exhibit when in combination with one another. . . . Our aim is not merely to explain the complex in the simple, but also to discover more about the simple by studying the complex. . . . [Perhaps] organisation is not truly a [concept] of science . . . but a philosophical idea . . . part of a scheme into which all phenomena can be fitted."

See Woodger, J. H., footnote 31, on the need for relevant, clear, initial ideas of organization; and footnote 69 for one definition of organization.

[2] Lotka, A. J., Elements of Physical Biology, *Baltimore*, 1925.

[3] Rashevsky, N., Mathematical Biophysics, *Chicago*, 1938.

[4] Needham, J., Time: The Refreshing River, *London*, 1943. [Includes several essays in which concepts of organization and integrative level are examined.]

[5] Kron, G., Tensor Analysis of Networks, *New York*, 1939, pp. xv, 538 especially. *"The keynote of this volume is 'organization'"*. [Italics in original. A text concerned with analysis and synthesis of electro-dynamic systems.]

[6] Hussong, H. M., An Analysis of the Group Concept, Univ. of Pa., *Phila.*, 1931. [History and analysis of the organic analogy.]

Spencer, H., The Study of Sociology, *New York*, 1903, Vol. 1, Ch. 14.

Gerard's concept of the org as the generalized unit of organization is the most recent extension.[7]

The speculation that analogy, metaphorization and generalization have encouraged during the last decades is here seen as constructive. After a point it arouses the skepticism and resistance of those who wish to further the cause of precision and fundamental progress in the sciences. We share, but do not equate the achievement of this aim to the finding of a precise formulation for any quantitative aspects the concept and data of organization may prove to have.

The concept of organization does appear to have wide application and fundamental significance in the physical, biological and social sciences, if not metrical aspects. Yet it seems too easy to make a survey of physics and biophysics, chemistry and biochemistry, biology, psychology and sociology, showing that organized entities and the principle of organization are to be found at every "level of integration."[8,9] This leads easily to the conjecture that organization and energy are the final concepts of science, whence "all" that remains is to clarify and quantify these jointly. A skeptical but realistic view, for example that of Lotka, retains expedient usages of the language and ideas of organization in the domain of individual organisms but seeks to reduce these as rapidly as may be to primitive logical, mathematical and material-energetic concepts.

A crude effort at reduction might seek simply to reduce without remainder biosocial process, and human history in particular, to the molar energetics of the 19th century, somewhat on the models set by Brooks and Henry Adams.[10,11] Sweeping attempts on the Adams model, which were not so crude as this hypothetical reduction, have

---

[7] Biological Symposia VIII, Levels of Integration in Biological and Social Systems, *Lancaster*, 1942, pp. 67-87. [No mathematical discussion or evaluation of the prospects for mathematization of its central concept.]

[8] Needham, J., Integrative Levels: A Revaluation. In reference, footnote 4.

[9] Novikoff, A., The concept of integrative levels in biology, *Science*, 101; p. 209, 1945. [Critical review of the literature.] Continuity and discontinuity in evolution, *Science*, 102; pp. 405-406, 1945.

[10] Adams, Brooks, Law of Civilization and Decay, *New York*, 1894. [Reprint with preface by Charles Beard, *New York*, 1943. Traces the mutual influence of Brooks and Henry Adams; considers critically their ideas of history as a science.]

[11] Adams, Henry, Presidential Letter, *Am. Hist. Ass'n.*, Annual Reports, 1894; The Education of Henry Adams, *Boston*, 1918; *New York*, 1931; The Tendency of History, *New York*, 1918; Degradation of the Democratic Dogma, *New York*, 1919.

now given place to detailed attack on the ideas of biosocial organization and its evolution, of equilibria of order and patterned mixed-up-ness in living systems, and their relation to the concept of entropy.[12] The use of metaphor, so free in the past, is more guarded so that the relation of these concepts to each other and to the physical law governing the increase of entropy, sometimes called the law of energy-disorganization, is not pre-judged. It is here believed to be necessary to guard carefully against confusion of the energetic-mechanical idea of entropy, first metaphorized as "disorganization" about the turn of the century, and the functional-adaptive or means-end idea of biosocial organization. In view of the fact that both sorts of organization are fairly well understood to involve order and statistical probability concepts in some essential way, it may not prove vain to hope for a unification of the concepts in place of a metaphorization that merely confuses them.[13, 14]

A question that appears answerable short of this unification can now be framed. Can we define organization in general terms, in metrical terms wherever required, and formulate possible laws of trend with the time of functional organization in an energetically non-isolated biosocial system, in a form consistent with the trend of entropy in the energetically non-isolated substratum of all biosocial systems.

If we leave aside entropy measure, and the second law of thermodynamics, which restricts the trend of entropy in energetically isolated (closed) systems, the sciences now offer nothing that can, without the serious reservation indicated, be called a measure of organization, either functional or other, and a law of its evolution in time. A unification of physical and biosocial concepts would include a valid restatement of present metaphorization—"entropy measures disorganization of energy"—and a conjectured analogue, e. g., "organization measures structure-function differentiation and integration"—together with

---

[12] Needham, J., Evolution and thermodynamics, *Science and Society*, Fall 1942. Lotka, A. J., Evolution and thermodynamics, [Reply to Needham], *Science and Society*, Spring 1944.

[13] Butler, J. A. V., Life and the second law of thermodynamics, *Nature*, 153; p. 153, 1946. [Critical note.]

[14] Schrödinger, E., What is Life?, *New York*, 1946. (". . . the device by which an organism maintains itself stationary at a fairly high level of orderliness (= fairly low level of entropy) really consists in continually sucking orderliness (i. e., minus entropy) from the environment." Chapter VI, p. 75, "Order, Disorder and Entropy.")

mutually consistent statements of any valid trend laws. A functional concept, turning as it does upon reference to means and ends, cannot be reduced to a mechanical-structural concept without remainder. This supposition of irreducibility is, however, not equivalent to the proposition that proof of the existence of entities corresponding to organization functionally defined, is precluded by the established existence of those entities corresponding to mechanically defined "organization."

Seeming attempts by Henry and Brooks Adams to reduce biosocial process to 19th century physics still appear as a vagary to physical scientists, who have no particular interest in functional organization. To others, who are so interested but who cannot profess to be clear on what the laws of energy and entropy mean for life process, it is at any rate apparent that the neglect of an essential remainder must be fatal. Careful reading of the Adamses shows that they did not offer their formulations as reductions, let alone as successful, but invited other proposals from any who "know better." The challenge has been thought uninteresting or impossible.[15] Of literature directed to related problems there is by now a great volume. Nearly all of it omits directly to ask our primary question, whose answering is prerequisite to the asking of the more sweeping questions. To rephrase this briefly: is it now possible to define organization functionally, or more generally and metrically in part at least, so that its proper scope as a concept of the sciences, physical, biological and social, shall become clearer.

If the answer is affirmative, there must follow a clear prescription of necessary steps for the enumerative, extensive or intensive measurement of organization in a number of areas of science, but particularly the biosocial. Examples of application to social aggregates, to organized or hierarchical social animal groups now so-called, must be given to illustrate and support the claim that organization can be measured. When a method is offered, such measurement will be seen to correspond to no very simple concept of measurement except in special cases. Further investigation would then be needed to show whether, in what manner, and under what conditions, functional organization increases, decreases or remains stationary with the time. The question whether it "always" tends to become greater with the time in the terrestrial biosphere[16] can then be phrased precisely. Meanwhile the record of offhand speculation and answers to these questions provides many

[15] Beard, Charles, in reference, footnote 10.

[16] Lotka, A. J., footnote 12.

clues to procedure and we shall continue to be indebted to it. In the next section we shall review some of the pioneering work in more detail.

For the most recent, directly relevant and sharply divergent views of our aim, the reader is referred to contributed papers by Needham [17] and Lotka.[18] The former encourages, the latter is skeptical of an attempt like the present one. Of course, we must here part company with any who deem the problem unreal, and go further than those who declare themselves unable to find reasons why organization cannot be measured, for we offer one method to that end and illustrate it concretely.

Whether in this paper or in subsequent work, organization is shown to be a scientific concept that only required reformulation to make its proper scope clear, or an over-rated philosophical idea without the importance increasingly assigned to it, there should be no quarrel with attempts at needed reformulation and the invention of measures for quantitative aspects. Discussion could for awhile direct its energies to an improvement of definitions and the means of measurement. While the best course may in the end prove to have been to dismiss the aim as futile on principle, the common language of organization will remain embedded in daily life to trouble a hoped-for science of management.[19] If the idea of organization as measurable entity could well go the way of the heat-substance phlogiston, it might like physical action, once a crude insight, become a general mathematical concept and measure of magnitude. The author takes the latter alternative to be more likely—that functional organization has quantitative aspects, that these can be expressed metrically in a non-simple sense, and therefore treated mathematically. The possibility of unification of the ideas of functional and mechanical organization cannot be excluded at this point of the discussion. We here show that the former can be expressed in terms formally similar to those in which the latter is already embodied.

---

[17] Needham, J., On science and social change, *Science and Society*, Summer 1946.

[18] Lotka, A. J., The law of evolution as a maximal principle, HUMAN BIOLOGY, 17: pp. 167-194, 1945.

[19] Mooney, J. and Reiley, A., The Principles of Organization, *New York*, 1939. [Historical and analytical study of organization in the ecclesiastical, military, modern industrial and political senses. The authors frankly take the position of advocacy in which their intensive experience and identification has placed them.]

2. PIONEERING ATTEMPTS TO DEFINE ORGANIZATION IN QUANTITATIVE TERMS

In this paper we shall explore one means for expressing quantitatively certain essential aspects of *social* organization. These aspects may be described very briefly and provisionally as functional differentiation between individuals, persons, or sub-groups and integration of these to constitute the functional whole. We desire also to formulate the idea of integrative level, as distinguished from degree of integration within a level. We shall proceed as rapidly as possible to apply the method to various special problems of human and other social animal "group organization" now so-called,[20] for which statistical data are already available or could theoretically be obtained. Consideration of the problem of individual animal organization is believed to be possible in formally similar terms but we shall omit this.[21] The method originates in concrete social situations, is statistical, can be represented formally in a very general algebra of mathematical groups, and already shows promise of developing a rationale in terms of mathematical probability.[22] Its basis is thus either superficially, or fundamentally, similar to that upon which entropy measure was early given its statistical-probability interpretation.[23] Much remains to be clarified in this connection.

For about three decades entropy has been interpreted by some few physicists, and by others given to wider speculation, as measuring the "disorder" or "disorganization" of the energetic-material constituents of the physical aggregate.[24] It should be noted that this interpretation

---

[20] Allee, W. C., Group organization among vertebrates, [expository], *Science*, 95, pp. 289-293, 1942. Influence of estradiol on the social organization of flocks of hens, [Biochemical factor], *Endocrinology*, 27, pp. 87-94, 1940.

[21] Needham, J., Order and Life, *New Haven*, 1936, p. 96f. [Organizer hormones and probability-field concept.]

Lotka, A. J., reference, footnote 18, pp. 171-172. [Critical note on measurement of organization of animals.]

[22] Lienau, C. C., Discrete bivariate distribution in certain problems of statistical order, *Am. Jour. Hygiene*, May, 1941, Sec. A, pp. 65-85. [p. 84, Suggestions for a probability representation of social animal hierarchical groups and their evolution.]

[23] Vendryés, P., "Vie et Probabilité," *Paris*, 1942.

Kostitsin, V., Mathematical Biology, *London*, 1939.

[24] Eddington, A. S., The Nature of the Physical World, *New York*, 1928, p. 70, 105.

arose late, that it was introduced merely as an explanatory metaphorization which neither was nor is now required for successful treatment of physical problems. The original usage was guarded, but the bright promise thus held out of filling what appears to be a fundamental need of the biological and social sciences—a measure of functional organization—has remained a promise.[25] We must therefore give attention not only to what is sound in this promise but also to the risk of overreaching the proper scope of the physical concept of "organization."

Entropy is the name originally given to a convenient mathematical construct corresponding to a one-dimensional additive physical magnitude. At first this was defined in terms of large scale or "molar" operations upon the aggregate which lead to measures of temperature, heat energy absorbed and mechanical work done by the aggregate. Entropy literally means "inturning," of the system's energy toward a condition of unavailability to external systems. Subsequently, as molecular concepts developed and were validated experimentally, entropy was reinterpreted as a measure of the state of internal "shuffling," "mixed-up-ness" or "randomness" of the molecular constituents of the aggregate. This condition could be expressed statistically in terms of the probabilities of configuration and motion of hypothecated constituents. The vast number of these elementary constituents, and the more or less correlated, more or less coordinated character, of their position and motion encouraged too-free use of animistic social terms such as swarm, disorder, anarchy, degradation, order, and organization.[26] Thus was the analogy drawn between the mechanical part-to-part and part-whole relationships, and those forms of functional unity and disunity apparent in swarms, hordes, or groups of social animals.[27] But if biosocial aggregates have usually been clearly seen to be integrated by individual and social communication and transport mechanisms, by learned responses to regular patterns of language, by personal affinities and consciousness of complex cultural values or purposes (unlike molecules except in poetry), the bare metrical and order relations of social groups remain fascinating to the mathematically inclined.

Short steps lead to the large conjecture that functional organization might be reducible to primitive metrical order and energy relations

---

[25] Schrödinger, E., footnote 14, the most recent of physicists to consider the relation of entropy to biological order.

[26] Adams, Henry, footnote 10.

[27] Eddington, A. S., footnote 24.

that are sufficient in the case of entropy.[28] But if this is mistaken, as we contend, the favor in which we continue to hold such parallels as the mechanically "organized" material particle aggregate and the corps of soldiers executing functionally organized maneuver becomes intelligible. It has proved similarly difficult to resist giving off-hand answers to the ultimate and urgent questions that arise as to the relation of chance and choice, of anarchic to free and to orderly social action, and the meaning and possibility of coexistence of determination and freedom in mechanically and functionally conceived systems.

Under the speculative pressure so generated it was natural that entropy should come to be spoken of free of qualification as if it were the sole possible and universal measure of disorder or disorganization. It was and, if this paper fails its aim, may remain all we have of possibly manifold measures of organization, special or general, funtional or other. The accepted form of the law of entropy change with the time claims to the highest degree of universality in application to energetically closed systems. Upon the removal of those warning qualifications that we have symbolized by quotation marks, speculation in these terms could and did proceed to the field of the social sciences with little restraint. That the essays of Brooks and Henry Adams should be entitled "The Law of Civilization and Decay," "Degradation of the Democratic Dogma," "The Tendency of History," "The Phase Rule of History," is indicative of the lengths to which the ideas of thermodynamics were stretched. The second law of thermodynamics, concerning change with the time ("evolution") of entropy in energetically isolated material systems has encouraged speculation as to the "evolution of organization" of a closed universe, and of (functional) organization in the energetically open and evolving terrestrial sphere of life.[29]

Whatever the final verdict on all this, it has enabled us here better to phrase our question as to the possibility of measuring biological and social organization in the functional sense. A quantification at the functional level can, we shall hope, be made independently and so as not to be in contradiction with the laws of the physical substratum of life, properly understood. A synthesis in which the mechanical and functional concepts are unified may thus be put down as a permissible

---

[28] Schrödinger, E., footnote 14, should be read with this reservation in mind.

[29] Whittaker, E. T., The Beginning and End of the World, *Oxford*, 1942, p. 40. [Riddell Lecture.]

hope. By this course we might conceivably reach mathematical expressions of various verbalized conjectures concerning the local or general evolution of functional organization. The modes of measurement of mechanical and functional organization being consistently prescribed, we might then seek historical or experimental verification in large-scale biophysical and social systems. That the maximum program is ambitious, if not illusory, this formulation is intended to show. But stated in this way it would seem to have some chance of quantitative empirical test.

So far as the author knows, no quantification of functional organization as such has yet been attempted.[30] As remarked above, doubt has been expressed that organization is a true concept of science. Yet in the same place we find the assertion that organization varies quantitatively; i. e., in intensity or extensive magnitude, and the further implication that it is a verifiable entity. The words "not yet" inserted after "organization is" would make the expressed doubt more conformable to the assertion of measurable entity. It is at any rate agreed among the keenest critics of its use that the idea of organization still lacks that clear initial formulation necessary to its use for theoretical purposes.[31] Meanwhile, definitions of functional organization continue to be expressed in familiar verbalized form. The recent *Dictionary of Sociology* (1944) does not raise the question whether a quantitative rendering of its own definition is possible, desirable, or fundamentally necessary. But the literature records a growing effort to grasp the suggestion of measurable variance contained in the idea of functional organization (e. g., Needham, Waddington), and it is recognized that the metaphorization of the mechanical concept of energy and entropy has been in some sense a promising sign-post to this end, and to ultimate unification.

Animated by promise of large reward was the early speculation that presented us with social energy and social entropy. But these concepts have not, under these names at least, materialized by standards of measurement, rationale and success that attach to their scientific ancestry. To the Adamses such hybrid terms were signs of a "science of history" to come. By 1944, five decades first allowed as the period of

---

[30] That is, in the sense of section 12, this paper.

[31] Woodger, J. H., Biological Principles, *London*, 1928. [Especially pp. 288-315, where the concepts of organization and integrative level are examined], and Waddington, footnote 1.

its emergence, had passed.[32] During this period, biometry,[33] physical biology,[34] biophysics,[35] econometrics,[36] psychometrics [37] and sociometry,[38] each with its mathematical "rationale," have commanded enthusiastic, reluctant and even derisive attention. They do not yet add up to historiometry, which is unknown as a science or a hopeful name for one. Charles Beard's preface to the 1943 edition of Brooks' "Law of Civilization" and an earlier essay states his grounds for regarding the Adams hope as chimerical. Social energy plays no serious part in this rise of the biosocial sciences. Social entropy, described as a "doctrine" (not to say a curious one) in the *Dictionary of Sociology*,[39] is not yet the name of one measure of functional disintegration among several recently invented measures.

The speculative pioneers would, however, take comfort from the increasing theoretical consideration of levels of integration and organization in natural systems, that has marked the last ten years.[40] Most significant is the fact that in 1940 R. C. Angell constructed our first empirical measures of the social integration and disintegration of American cities, using statistics of crime, broken marriages and welfare effort in relation to population mobility, heterogeneity and other relevant socio-economic factors.[41] The half-century since 1894 would probably not have disappointed the Adamses.

Contributors to the present collection of essays take up various positions on the possibility and need for measurements of functional organization. Lotka's skepticism is clearly stated in his recent work [42] but does not preclude an attempt like the present. He has privately

---

[32] Adams, Henry, Presidential Letter, footnote 11.

[33] Biometrika, 1901- (Journal).

[34] Lotka, A. J., footnote 2.

[35] Rashevsky, N., footnote 3.

[36] Econometrica, 1932- (Journal).

[37] Psychometrika, 1936- (Journal).

[38] Sociometry, 1937- (Journal).

[39] Social entropy: "The doctrine of inevitable social decline and degeneration derived from the field of physics and particularly from the Second Law of Thermodynamics, which conceived that the energy of the universe, fixed and limited, was being given off to futile heating of empty space, and was not being replaced," Dictionary of Sociology, *New York*, 1944.

[40] Footnote 9.

[41] Angell, R. C., The social integration of American cities of more than 100,000 population, *Amer. Soc. Rev.*, 12, pp. 335-342, 1947.

[42] Footnote 18.

expressed the opinion that direct attack on the organization problem is not essential to progress of the life sciences, that we may continue to look for its solution as a logical incident in the development of postulates making no specific mention of organization. An earlier essay of Needham's declared that he could find no reason why organization cannot be measured and implies that fundamental progress may be served by an effort to effect measurement. The position taken here is that if it be possible now to quantify and empirically to measure the mechanical substratum concept entropy ("disorganization"), and also the biosocial-functional concept of organization in a consistent way, if a unification of the two concepts is possible, adherents of these views should give an early demonstration.

If clear initial formulation of the idea of functional organization is prerequisite to its expression in quantitative terms, it has only recently become evident that such clarity has yet to be attained even in those areas of science where it has been considered to play a legitimate part in non-quantitative theory. Clarity and quantification may go hand in hand. The earliest critical and constructive essays in which the needed clarification is sought but in which the question of specific means for quantification is left unanswered as subsidiary, are due to Hussong [43] and Woodger.[44] Whether separable or not, the questions are of equal recency.

In the *Journal of Sociometry* (1937-), if anywhere, should the concept of organization as it applies specifically to the social group at last have received its own mathematical tretament. The motivation of the papers appearing in this journal has, however, been in the direction of psycho-social therapy ("sociatry") rather than the finding of mathematical social rationale and statistical method as such.[45] But formal priority for the application to the social group of mathematical statistics and matrix algebra does belong to contributors to this journal. Of special interest to us is the group interrelation matrix. This, with important modifications in rank order notation and more careful probability interpretation of the matrix elements, is used in the present method and, properly speaking, is a mathematical representation of an

---

[43] Footnote 6.

[44] Footnote 31.

[45] Dodd, S. C., Sociometry delimited, *Sociometry*, 4, pp. 200-205, 1943. [The interrelation matrix, sociometry vs. sociatry.]

essential aspect of social organization.[46] But it has been left indefinite with respect to the primary question of our objective. It constitutes "measurement of organization" in too loose a sense to shed light upon the question of the relation of mechanical and functional organization.

A recent work by the mathematician von Neumann and the economist Morgenstern is entitled "The Theory of Games and Economic Behavior." [47] The title does not at once reveal its interest in the present connection. The work deals mathematically with the relations between two or more persons or alliances in play or economic activity. That it must involve some formalization of the social group concept in probability terms is apparent. Therefore the idea of functional organization may come in for careful consideration, if not be given mathematical representation. The terms social organization are in fact used frequently throughout the volume. In one place it is even said that a treatment of sports statistics in probability terms is inadequate as a basis for a theory of social organization. This cannot be disputed, but it carries the implication that the work should have been entitled "Theory of Social Organization." Still no general definition of social organization in the functional sense is offered, nor does it emerge from the work considered as a whole.

Sociologists who would be attracted to such a title are entitled to expect at least two of the fundamental problems of social organization to be treated, however excessive the demand may appear at the outset. One demand is for at least one representation of social status or hierarchical groups in the course of temporal evolution. The other is the problem of representing (organized) groups as distinguished from "mere" aggregations. Thus the spectators at a fire, and the fire company with their chief, differ in a way that leads us to call the latter a group and the former a mere aggregate or, at best, a trivial group. Current expressions of this difference involve reference to function (watching or extinguishing a fire) and may use the term organization. Differentiation of function among persons or sub-groups, integration of these to form the functional whole,—however we choose to resolve the subsidiary concepts of functional organization or to represent complex functional groups (e. g., the group firemen-spectators)—the clari-

[46] Moreno, J. L., Contributions of sociometry to research methodology in sociology, *Am. Soc. Rev.* 12, pp. 287-292, 1947.

[47] von Neumann, J. and Morgenstern, O., Theory of Games and Economic Behavior, *Princeton*, 1944. [Chapter I, and organization passim.]

fication, mathematical presentation, and empirical measurement of quantitative aspects of organization are our problem.

The author has previously offered one class of means for the representation of a lower-animal hierarchical "social" group in a state of temporal evolution.[48] The problem of representing two co-existent or associated groups whose type or degree of organization differs is logically related, and at this point the problem is conjectured to fall within the scope of the same method, if this be conceived generally. Since the method grows out of consideration of statistical records of social animal behavior, it does not stop with formal representation for lack or relevant present data. For the same reason it finds application to statistics of organized sports. Such applications are illustrated in this paper without implying that they form an adequate basis for a theory of social organization or exhaust the content of the concept.

3. GROUPS AND SOME OF THEIR ORDER RELATIONS

We must distinguish between an aggregation of people or, more generally, of social animals, and a social group[49] of such units. To express the distinction in metrical terms is part of our task.[50] A group is an aggregate of two or more social animals characterized by a persistent or time-invariant pattern of interrelations. The isolated unit is an important abstraction whose consideration we mainly neglect. "Temporary" groups are limiting cases whose inclusion depends upon the unit of the time scale.

In the case of human groups the pattern is usually held to be psychological, to involve conscious purpose, and the interrelations in which a time-invariant pattern can be found are not necessarily limited to those which are objective (structure and motions of the system actually or potentially observable by at least two persons). A human group is ordinarily held to be self-recognized as such; i. e., that it is not a mere aggregate and it may be recognized as such by the members of other groups in virtue of the established and perceived patterns of internal and external relations. Limiting sorts of patterned relations

[48] Footnote 22.

[49] *Encyclopaedia of the Social Sciences*, Vol. 7, p. 178. Any [human social] group is constituted by the fact that there is some interest which holds its members together.

[50] In section 11 this is undertaken in terms of a measure of integration.

of man to other animals suggest the need in general sociology for a clear concept of interspecies or complex social integrations. In another direction the concept must encompass the club, the league, and the high council or court.[51]

In the case of lower animals and insects called social, objective patterns of interrelation having greater or less temporal stability have been verified in natural and laboratory aggregations.[52] That there is involved conscious mutual recognition by the individual animals may be the naive observer's assumption or projection. Pattern is, however, a complex inference based not on reports of introspection as it might be in human social studies, but upon a statistical record of behavior in staged or unstaged contact interactions between animals. From such records a large variety of time-invariant order-statistical representations of group relation can now be derived. The order relations currently chosen as the analytical elements of organization in animal groups are known under the names pecking order, dominance-subordination, hierarchical structure and, more generally and vaguely, as "group organization." Organization is treated as understood or to be understood partly or fully in terms of the invariances selected for study. Such attenuated descriptions are thus of restricted, if essential, aspects of the complexity that "group organization" of social animals presents to the observer at the outset. Unlike that of the theoretically neutral observer of a physical aggregate, the regulative and anthropocentric positions of the primary observer must be taken into explicit account,—that is, the primary human experimenter and "his" animals form an aggregate or a group of a sort, if not a social group in the narrower sense of human grouping. At any rate it is recognized that we must distinguish factors according to their relevance and origin in the wider situation, however naive the primary experimenter may have been. We here shall agree that our concept of the social group and its representation must be comprehensive enough to include limiting sorts of aggregates of two or more members of naturally symbiotic or artificially cooperating species. If

---

[51] Social Organization, *Encyclopaedia of the Social Sciences,* Vol. 14, pp. 138-141. "Every human group is organized; its individual components do not behave independently of one another but are linked by bonds, the nature of which determine the social unit. Kinship, age, sex, co-residence, matrimonial status, community of religious interests are among the unifying agencies; and in stratified societies members of the same level form a definite class".

[52] Allee, W. C., footnote 20.

we do not provide for such cases we may be at a loss as to how to represent simple and complex interspecies aggregates, of which the horseman, a classical symbol of cooperation and contest, forms a constructive example for us.

The concept of the *mathematical* group is simpler and better defined than the social group in that it involves initially, at least, as few as we choose of logical or mathematical "laws" or "relations" — between "members" of a set of two or more simple mathematical "classes" of elements logically associated to form a "group." The elements and groups represented may be complex mathematical elements or even social groups in turn. We might therefore expect the mathematical group concept to be useful in the partial definition of the social group and its organization, though the elements and relations of social groups are not in general well understood or simply represented. This expectation is correct. Here particularly social terminology and partial analogy (class, group, relation, association) will force themselves upon our attention despite the clear mnemonic intent of the verbalisms employed in mathematics.

One of the simplest groups of immediate use to us is that of the permutations $(m + 1)$ at a time on a set of $(m + 1)$ "things," letters or numbers, say the "natural" numbers to m (m finite), including zero by extension; $j = 0, 1, 2, 3 \ldots m$. These permutations form a set of finite groups containing $(m + 1)\ m\ (m - 1) \ldots 3 \cdot 2 \cdot 1 = (m + 1)!$ members. In general, $m = 1, 2, 3 \ldots$ with or without limit. Any member of this set of groups can be formed, representatively, by substituting progressively one for another of the members of the aggregate $j = 0, 1, \ldots m$. For example, the members of the first three groups can be represented (0), (01, 10), (012, 102, 021, 120, 201, 210) for $m = 0, 1, 2$. The operation employed in passing from one member of the group to another is of the character known as "associative" in the strict mathematical sense. It is non-commutative and each of the groups includes unique inverses corresponding to its members such that the operation it represents, performed upon that member, reproduces the unit or identical member of the group, here denoted 0, 01, 012, etc. The "group members" are "relational" entities in the sense that they have "properties not belonging to their constituents" but to the "group" only. Our tendency toward the social metaphorization of technical language may tempt us here to go the whole way and use the term group "organization." But despite the widespread and

growing incidence of this usage in science and technology, mathematicians have not succumbed to it.[53, 54]

The integers $r = 1, 2, \ldots$ have been used almost universally to identify and rank ordinally the members of physical or biosocial aggregates. With due care we can extend this usage to biosocial groups and their parts. In the latter case m is one at least; $m = 1, 2 \ldots$. This specification of the range of m may be written $m \neq 0$, thus stating that a non-trivial group contains at least two individuals or that the cardinal number of the group is 2 at least, the upper limit being specified only as finite. The excluded case of the isolated part, unit, individual or group, $m = 0$, is clearly of fundamental interest to abstract theory. It is without exact empirical counterpart unless the sole contender, the lone wolf, the hermit, the castaway or terrestrial man are viewed as isolates.

The use of the extended integer notation $r - 1 = j = 0, 1 \ldots m$ (that is, the traditional range extended to include zero) has a number of clear advantages, but also certain disadvantages for our purposes, and we shall now point out some of these to guard against an uncritical, or too-restricted, representation of social groups in general.

The members of a social group are always assumed to be homogeneous in at least one respect, which is understood or is not always made fully explicit, and to be differentiated in at least one respect relevant to the difference between a mere aggregate and a social group. "At least one respect" implies the possibility of indefinitely many (m) and the impossibility of exhaustive description except in theory. The integers j may be used to designate occupancy by the individual or the members of a sub-group collectively, of one of a logically complete set of mutually exclusive physical, biological, psychological or social categories. These categories may be dichotomous, qualitative and metrically ungraded as, for example, sex (designated 0, 1). They may be ranked and manifold, as dual or multiple hierarchies of social status, military ranks, or ecclesiastical grades; for example, nobles and commoners, officers and men, priests and parishioners, (similarly designated 0, 1). They may form an ordered but unmetricized selection of some continuous infinite series of intensive or extensive magnitudes, as temperature or extension.

In case of individuals ranked in ordinal sequence, the integer $j = 0$,

---

[53] Footnote 5.

[54] Footnote 47.

1 . . . m, increased by unity, is used to denote "rank" or ordinal number in the aggregate of differentiated individuals counting from the lowest or highest in the scale and assuming no two alike in magnitude. By abandoning the old convention r for the extended integer designation $r - 1 = j = 0, 1 \ldots$, we specify at once the cardinal number of members of the aggregate "other-than-j" and thus leading or superordinate to "j," while by $(m - j)$ we specify the like number of the others following or subordinate to "j." Thus "$j = 0$" has *no* leader; i. e., is preceded by *none* of the other members of the aggregate, "$j = 1$" by one other, etc., while m is preceded by m others in the scale. Similarly, $(m - j)$ is the cardinal number of followers of "j," members of the aggregate of individuals "other-than-j" that follow "j" in the scale. Thus, j serves as identifying label, as ordinal sign, and as a cardinal number specification.

In number theory $j + 1$ has been called the "successor" of j, which is the "ancestor" of $(j + 1)$. Here in the logic of mathematics we come upon a metaphor derived from primitive genetic and social experience. It serves as a mnemonic for a familiar relation of succession.

In the case of a formally constituted human organization[55] characterized by a ranked sequence or line of grades, e. g., the military, ecclesiastical or industrial hierarchy, the number of individuals or sub-groups that are formally subordinate to some individual or sub-group j and of the next-lower grade is ordinarily greater than unity. Here a multi-dimensioned, reticulate, or echelon representation is called for. If an aggregate of coordinate equals functions more or less as a unit, it may be treated as if it were integral, a bloc or group in the chain of command, "the line" as it is called. The number of "lines" originating from one organizational center or head may be very great. In the formally similar case of two or more contestants finishing a race abreast or tied, the coincidence at the finish cannot be described fully in terms of a one-dimensional order. We may then choose conventionally to linearize the representation of such multi-dimensional aggregates and represent them simply by $j = 0, 1, \ldots m$, enlisting the aid of random numbers, coin toss or photo-finish mechanisms to resolve ties for the same rank.

Our notation for the so-called "rank order" of a set of $(m + 1)$ elements (unlike the traditional notation $r = j + 1$) takes into reckon-

[55] Footnote 19.

ing the finding of Kendall and Smith that initial neglect of the distinction between ordinal and cardinal numbers has been a source of difficulty in the theory of rank-order statistics.[56]

In cases where a linear representation of $(m+1)$ elements applies logically, the ratio j/m can be interpreted as the probability that a member of the aggregate "other-than-j" is super-ordinate to "j." Similarly, $(m-j)/m$ is the complementary probability that one of the m other members is subordinate to "j" in the order represented by $j = 0, 1 \ldots m$. Generalization to bivariate and multivariate systems is not difficult.

Integer notation has always had the initially unclear disadvantage that the attention of its users tends to become directed upon cases of conveniently one-dimensional social order, to the neglect of the multi-dimensional and non-transitive[57] social order relations that are known to be commonplace realities for $m = 2$ or more. These problems do not arise in the fundamental case $m = 1$, which we shall call the dual group. They must be faced in small multiple groups, indeed, when $m = 2$. Otherwise we must logically confine our interest to a very restricted field of social reality.

4. MULTIPLE GROUP AS A DUAL GROUP

A race is a contest of speed for material stakes such as money, or non-material stakes such as social status, between at least two contestants or complex groups of contestants. The foot-race and the horse-race are perhaps the most familiar examples of cooperative-contest in which many, not all, of the quantifiable aspects of higher forms of "group organization" appear. A group of two horsemen in competition is a constructive illustration for us and a popular symbol for a large and much more complex class of dual social groups. These consist of two dissimilar individuals or sub-groups, possibly of distinct biological species as horses and men, engaged in cooperative contest or mortal conflict with similar groups, and differentiated in one or more characteristics or ability relevant to success in contest. The file, phalanx or squad of military horsemen, while scarcely a simple extension from the group of $(m + 1)$ racing horsemen, aids us in extending the notion of the dual group to complex and multiple groups

[56] Footnote 22.
[57] Footnote 47.

in which at least two superposed "levels of integration" are exhibited. The units of these levels are commonly thought of in an order of complexity: horses, men; horseman; squad; etc. The value of the race as a model for a beginning theory of contest in which status or other biosocial values are at stake is enhanced by the fact that here our current working concepts of probability, gaming, and accurate voluminous distributional statistics of the race are found in convenient correlation. They are in fact very often set forth in what we here call the dual form of representation.

The representation we shall use for the race is formally similar to that now used in presenting the statistics of other branches of human sport, in which dual contest between individuals, teams or groups leads to the division of points or stakes and to statistical records of the division. The representations of "organized" groups of social animals derived from laboratory studies have a similarity to these which is not surprising. Indeed the range of choice of representations is narrow, given our rather elementary stock of socially shared ideas of number and recording technology. The apparent universality of certain statistical forms is thus accounted for partly by the primitive form of our thought about aggregates and groups generally, and social groups in particular. Indeed one-dimensional and dual forms are attractive to us partly because of their supposed logical or psychological primacy which is signalized by the terms "natural" or "simple." These forms must therefore be considered fully, to define the advantages and disadvantages of modes of thinking that could become more widespread than is justified by their adequacy for social theory. We will thus be forewarned in specific ways and generally, that an over-simple representation can become an unseen obstacle to progress and quantification. The long overlooked and apparently trivial matter of representing rank order so informs us.

The following example is drawn from American statistics of racing.[58] One hundred chronologically successive horse-races staged in the United States during June 1942 were selected to be alike in that for each race eight entries started, and finished without ties. The entries of each race were ranked by starting odds and finishing order. A very few paired entries equal in starting odds were separated and ranked by post position, which is assigned at random to distribute its presumptive advantage. The two chronological sub-series of 50 races

[58] *Racing Form Monthly*, June 1942, New York.

each, had the distributions of public favor and outcome shown in Tables 1 and 2.

It is seen that the favorites won very nearly one race in three. For other values of m the results are similar, $m = 7$ having been selected

TABLE 1

Starting odds and rank at finish
First series: 50 consecutive races, June 1942

CONTESTANT RANK BY STARTING ODDS	RANK AT FINISH	
	Won	Other ranks
Favorite	16	34
Other ranks	34	..

($m + 1 = 8$ contestants per race; $N = 50$ races)

TABLE 2

Second series: 50 consecutive races, June 1942

CONTESTANT RANK BY STARTING ODDS	RANK AT FINISH	
	Won	Other ranks
Favorite	17	33
Other ranks	33	..

($m + 1 = 8$ contestants per race; $N = 50$ races)

as fairly representative of United States racing. The ratio of favorites-won to favorites in other positions is close to the rule that "fair odds" on a favorite are 2 to 1 against winning. The favorite is here and ordinarily defined as that contestant whose pari-mutuel "starting price" is the smallest; that is, the ratio of dollars laid that other entries will win to dollars laid that the favorite will win is smallest or "shortest" at the moment of the start. More generally, odds are quoted without reference to monetary unit or actual distribution of dollar bets.

Numerous basic data on racing are matters of precise, uniform and regular public record. They include the rank orders of the contestants at lines along the course and at the finish expressed as permutations of post position, $r = 1, 2, \ldots$, with other current and historical particulars of the race, horses, riders and owners. If our 100 races had actually been truly dual instead of multiple, the unspecified entries would be determined as 16 and 17 races by the condition that each array must have a total equal to 100, the number of races represented by the table. Thus, although $m = 7$ and the actual detail is extensive, the end results are here recorded as if they originated from 100 races, $m = 1$. This illustrates the statistical contraction of the case $m > 1$ to $m = 1$, corresponding to a familiar social logic by which multiple social groups are often contracted or "grouped" and treated as if they were dual groups. If the aim of this contraction is simplicity the process is not simple.

5. MULTIPLE GROUP

We now proceed to an example of the case $m = 7$ based upon the detailed data of the preceding example, where the seven other contestants were first treated as if undifferentiated; i. e., as if $m = 1$. The $2 \times 50 \times 8$ contestants are now distinguished according to $8 \times 8 = 64$ classes by the starting odds order of public favoritism and by order of finishing. (See Tables 3-7.)

TABLE 3

Distribution of starting odds order (j) versus finishing order (k).
$m + 1 = 8$ contestants; $N = 50$ races per array.
Sample no. 1; 50 consecutive races U. S., June 1942
(Tabulated from Racing Form Monthly, June 1942)

	k							
j	0	1	2	3	4	5	6	7
0	16	10	9	6	6	2	0	1
1	8	10	8	10	6	1	6	1
2	10	9	8	12	7	2	1	1
3	6	11	6	5	6	4	7	5
4	5	3	10	3	4	15	7	3
5	4	6	5	8	8	7	7	5
6	0	1	4	4	6	12	13	10
7	1	0	0	2	7	7	9	24

It was found by tabulation and calculation, for which space is lacking here, that there was a remarkably close relationship between odds corresponding to the geometric mean starting price P_j of each favoritism order of contestant and the relative frequency of appearance of these orders of contestants in each finishing order (Table 7). There are also seen diagonal and central types of mathematical symmetry within each of the Tables 3-5 as well as close resemblances between them. These points are brought to attention briefly to illustrate the relations between

TABLE 4

Sample no. 2; 50 consecutive races U. S., June 1942

	k							
j	0	1	2	3	4	5	6	7
0	17	13	4	5	5	5	1	0
1	13	10	6	5	10	2	3	1
2	6	7	10	9	7	6	1	4
3	4	7	14	9	5	3	4	4
4	4	5	6	5	7	7	9	7
5	5	2	5	8	7	9	8	6
6	1	3	3	3	6	13	12	9
7	0	3	2	6	3	5	12	19

TABLE 5

Sum of samples no. 1, no. 2
$m + 1 = 8$; $N = 100$ races per array

	k							
j	0	1	2	3	4	5	6	7
0	33	23	13	11	11	7	1	1
1	21	20	14	15	16	3	9	2
2	16	16	18	21	14	8	2	5
3	10	18	20	14	11	7	11	9
4	9	8	16	8	11	22	16	10
5	9	8	10	16	15	16	15	11
6	1	4	7	7	12	25	25	19
7	1	3	2	8	10	12	21	43

TABLE 6

Correlation of starting odds and starting odds order for sum of samples no. 1, no. 2

	STARTING ODDS									
	Dollars against per dollar for j to win									
j	.5<1	1<2	2<4	4<8	8<16	16<32	32<64	64<128	128+	
0	5	55	37	3	..	..	..	..	..	100
1	..	5	77	18	..	..	..	..	..	100
2	..	..	25	74	1	..	..	..	..	100
3	..	..	1	68	31	..	..	..	..	100
4	..	..	..	33	51	16	..	..	..	100
5	..	..	3	10	45	33	9	..	..	100
6	..	..	..	..	15	45	36	4	..	100
7	..	..	..	..	1	21	40	37	1	100
	5	60	143	206	144	115	85	41	1	800

TABLE 7

Geometric mean starting odds versus starting odds order j. Expected and actual frequency of j-won; sum of samples no. 1, no. 2

ODDS ORDER	G.M. ODDS : 1	* EXP. % WON	RACES WON
j	P_j	$\frac{1-\hat{p}}{1+P_j} \times 100$	$n_j(0)$
0	1.84	31.9	33
1	3.10	22.1	21
2	4.79	15.6	16
3	6.96	11.3	10
4	10.06	8.2	9
5	14.42	5.9	9
6	27.66	3.2	1
7	50.21	1.8	1
Total		100.0*	100

* $\sum_j \left(\frac{1-\hat{p}}{1+P_j}\right) = 1$

$\hat{p} = .0942$

the underlying subjective (individual) estimates of probability, the public consensus, and their objective expression shown in the betting and averages of the outcome frequencies. We remark most particularly upon the possible role of mathematical probability theory in analysis of mathematical symmetry and other regularities in statistics of contest for social rank. The first of these possibilities to emerge is examined later. Here we limit the discussion to those essentials that lead to the matrix algebraic representation for the mathematical permutation group implied by such a description of contest for rank. Some other mathematical and statistical particulars of this representation were developed in an earlier paper.[59]

6. GROUP ALGEBRA ($m = 1$)

Restricting the discussion to $m = 1$, let I_0 denote the unit matrix of order $m + 1 = 2$. I_0 is defined as square table of units and zeros written thus:

$$I_0 = \begin{array}{c} \\ j \downarrow \end{array} \begin{array}{c} k \\ \longrightarrow \\ 0 \quad 1 \end{array} \left\| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\|$$

in which the row and column designations are $j = 0, 1$; $k = 0, 1$. We abandon at the outset an established convention in mathematics which is to number the arrays $r = j + 1$. By interchanging rows or columns once we obtain I_1, the representative matrix of the sole other member of the permutation group corresponding to $m = 1$:

$$I_1 = \begin{array}{c} \\ j \downarrow \end{array} \begin{array}{c} k \\ \longrightarrow \\ 0 \quad 1 \end{array} \left\| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right\| .$$

I_0, I_1 thus represent algebraically the members of the permutation group of two elements, "things" or sub-groups (denoted 0, 1) taken two at

[59] Footnote 22.

a time. Multiplication of I_0 and I_1 in the matrix sense corresponds to, and represents algebraically, the operation of permuting 0 1 and 1 0. Addition of matrices I_0 and I_1, in weighted proportions, has an immediate interpretation. Restricting (n_0, n_1) to positive integer values, including zero:

$$n_0 I_0 = \left\| \begin{matrix} n_1 & 0 \\ 0 & n_1 \end{matrix} \right\| \qquad n_1 I_1 = \left\| \begin{matrix} 0 & n_1 \\ n_1 & 0 \end{matrix} \right\|$$

$$n_0 = 0, 1, \ldots \qquad\qquad n_1 = 0, 1, \ldots$$

and adding

$$n_0 I_0 + n_1 I_1 = \left\| \begin{matrix} n_0 & n_1 \\ n_1 & n_0 \end{matrix} \right\|$$

or writing

$$n_0 + n_1 = N \quad \text{where } N = 1, \ldots$$

we have

$$\left\| \begin{matrix} n_0 & N - n_0 \\ N - n_0 & N_0 \end{matrix} \right\| .$$

This representation combines the manipulative convenience and power of matrix algebra and the special advantages of our notation $j = 0, 1, \ldots$ in the cause of statistical analysis of social groups. The step to a probability-measure interpretation of statistical ratios like n/N is at once suggested, but must be made with caution.

According to this representation, the matrices of Tables 1 and 2 are added algebraically to form the representation of the aggregate of 100 races. Assuming $N = 100$ and supplying the unspecified terms:

$$(1) \text{ and } (2) = \begin{matrix} & 0 & 1 & \\ 0 & 33 & 67 & m + 1 = 2 \text{ contestants} \\ 1 & 67 & 33 & N = 100 \text{ races} \end{matrix}$$

where $N = (16 + 34) + (17 + 33) = 100$ races.

It is seen that under certain conditions of contest and selection of data the relative frequencies of the joint classes (of favoritism and outcome) may, as N increases, tend statistically toward a number characteristic of the contestant population and conditions of contest. Taken as a general postulate this leads to empirical formulation of statistical rules of experience such as "two-to-one against favorites."

These can, theoretically, have as high a degree of numerical precision and empirical validity as any of the well-known biostatistical or actuarial statements of regularity concerning human or other aggregates. Whether such invariances are sufficient to support an organized gambling sport, a life insurance business, or large scale social planning are questions of fact. The reader will form his own impressionistic or technical sampling estimates of statistical homogeneity exhibited by the above examples. Questions of sampling, of biosocial homogeneity, of variation and causation are beyond the scope of a discussion concerned with formal representation and with the specification of simple or complex types of distribution adequate to comprehend the present range of experience.

7. EXTENSION OF THE ALGEBRA TO MULTIPLE GROUPS ($m = 1, 2 \ldots$)

When $m = 2$ there are $(m + 1)! = 3 \cdot 2 \cdot 1 = 6$ representative matrices corresponding to the group. These can be generated by permuting rows or columns of I_0:

$$I_0 = \begin{array}{c} \\ j \\ \downarrow \\ \end{array} \begin{array}{c} k \rightarrow \\ \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \left\| \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right\| \end{array} \begin{array}{l} \\ 0 \\ 1 \\ 2 \end{array} .$$

We can denote the six group members $I_0, I_1 \ldots I_5$, and the corresponding integer coefficients by n_i, $i = 0, 1, \ldots 5$. Matrix multiplication is the corresponding algebraic operation by which one representative permutation matrix is derived from any other in the group, which in general has $(m + 1)!$ distinct members. Numerous theorems of matrix algebra are now enlisted in the analysis of such tables as Tables 3 and 4. For this case I_i contains $8! = 40{,}320$ members. Whence the need for a systematic algebra and a classification of the membership in relation to properties of the integer index i.

We now interpret addition of terms like $n_i I_i$. Thus, addition of the N favoritism-outcome races of all classes i is symbolized

$$n_j(k) = \sum_i n_i I_i \qquad i = 0, 1 \ldots (m + 1)! - 1.$$

This represents the result of classifying the separate races into the sorts I_i and adding matrix-wise or, equivalently, of tallying in one frame the frequencies $n_j(k)$ of all $(m+1)N$ contestants according to $(m+1)^2$ bivariate order classes $j\ k$. The matrix sum so obtained can be represented completely by an equivalent matrix having m^2 or fewer elements, m^2 being the number of elements not already determined by the condition that each array totals to N_0.

As we shall show, probability theory already provides at least one ideal matrix who entire set of $(m+1)^2$ elements is determined given the integer parameter m alone. In general, $(m+1)$ or fewer elements of the matrix are of practical interest, though more may be required fully to characterize the matrix sum. For example, we add Tables 3 and 4 to indicate the sort of sample to sample regularity and internal symmetry that emerges when $N = 100$. The interesting diagonal or central symmetry that already appeared for $N = 50$ suggested that there were in operation restrictive biosocial, general statistical and, of course, strictly formal relations. Statistical experience of rather widely varying content but formally described by the permutation group as we have here represented it, has this sort of central or diagonally concentrated symmetry. In this particular case the symmetry means that the consensus order of the starting odds and outcome order are, in the general statistical sense, "positively associated." In the more special terms of contest, the odds at the start bear a correct order relation, on the average, to the space order at the finish and conversely. Variations from the average order occur with their own kind of regularity. The formal restriction that each row or column has the sum N is obvious as a source of regularity but not in the mode of its operation. Probability theory throws interesting light on the symmetries of the table. It also gives us some simple distributions in terms of variables j, k and constant parameters like m.

8. A ROLE FOR PROBABILITY THEORY IN CONTEST

In the foregoing we have arrived at the concept of the race as a controlled contest for rank between the members of a multiple group of size $(m+1)$. It seems long to have been recognized, if only tacitly, that such contests involve what have been called "a priori" probabilities, of prospective performance, and experienced or "a posteriori" frequencies of performance, as these terms are used in the theory of inverse probability due to Bayes and others. To the author's knowl-

edge this observation has never been stated explicitly. At least no attempt has been made to connect relevant observations of social contest with the classical theory of inverse probability. The suggestion is apt to arouse a derisive skepticism growing out of a century of controversy and early misuse of the theorem of Bayes on inverse probability. Recent students are still divided on fundamental questions, but if such theory can play a part in contest problems, these same questions may come to be seen in a new light.

In 1940 the author was led by these considerations, developed deductively, to the discovery that a probability distribution in two discrete variables: $j = 0, 1, \ldots m$; $k = 0, 1, \ldots m$; already deduced by Laplace and elaborated by Karl Pearson on the basis of inverse probability, is a useful abstract model for rank order problems. This was in fact the first result of the shift from the rank notation $r = j + 1 = 1, 2 \ldots m + 1$ traditionally used in the statistical theory of ranks and grades. The hampering confusion between ordinal and cardinal number concepts was this cleared up and the way opened to this conclusion. The resulting useful interpretation of Laplace's distribution in the social context of contest for rank, and the organization of leader-follower groups generally, is touched upon here without entering upon the controversies that flowed from Laplace's original argument or the deduction of related more general distributions. We shall state the algebraic result and then interpret its terms in the race. First we describe the notation and define a general class of bivariate discrete distributions.

Consider first the probability $Z_j(k)$ that a contestant j among m others of specified prior rank $j = 0, 1, 2 \ldots m$ will be found in any rank $k = 0, 1, 2 \ldots m$, at the finish of the contest. It is clear that Z is a "bivariate" discrete probability function of the "variates" j and k, and of one or more constant parameters not exceeding m^2 in number, among which m (here interpreted as the cardinal number of other-group members) may or may not be found explicitly.

Thus the sum $n_i I_i$ divided by $\Sigma n_i = N$, is such a function, indeed a highly general one for our purposes. In particular, when $n_i =$ constant, Z is independent of j and k. All $(m + 1)^2$ elements are then equal to $1/(m + 1)$. This means that the possible outcomes are equally probable for all contestants. This distribution may be named the null. At the other extreme, of inequality of chances, if $n_0 = N$, every one of the N contests shows the contestant j to be in outcome rank $j = k$.

That is, the prior rank j exactly and invariably predicts the outcome to be j. This distribution may be named the unit.

We may now wish to construct or deduce from elementary considerations of probability some intermediate types of Z. On intermediate models of distribution leading favorites would tend to lead at the finish, the least favored tend to trail, while moderately favored contestants would tend to deliver a middling performance. Some models should be such that no prior class is to be deprived of all chance of winning, while the favorite may well run last. In statistical theory, as in other branches of applied mathematics, we have preference for formulae with very few parameters. Z, as defined above, has $(m+1)!$ — 1 or fewer independent frequencies, since there are at most that many linearly independent frequencies n_i. Originally the author conceived the search for intermediate distributions in these terms of form and parametric simplicity.

The earliest result was the one-parameter Laplace form of Z:

$$Z_j(k) = \frac{m+1}{2m+1}\binom{m}{k}\binom{m}{j} \Big/ \binom{2m}{j+k} \quad \begin{matrix} (j, k = 0, 1, \ldots m; \\ m = 0, 1, \ldots) \end{matrix}$$

where

$$\binom{m}{k} = \frac{(m!)}{(m-k)!\,k!}$$

It is centrally symmetric, completely determined when m is specified, and is intermediate in the above sense.

In particular, when $m = 0$, the case of sole contestant, first rank $k = 0$, should go to that contestant with certainty. Here, in fact, $m = 0$, $Z_0(0) = 1$. A horse-race in which only one of $(m+1)$ scheduled contestants appears at starting time is conventionally awarded by default. The sole contestant formally establishes claim to the stake by a "walkover" of the finish line. Such races are recorded.

When $m = 1$,

$$Z_j(k) = \frac{1}{3}\begin{Vmatrix} 2 & 1 \\ 1 & 2 \end{Vmatrix}.$$

Except as a possible model for selected classes of contestants and specially controlled contest, this distribution will not be found applicable as it stands to present experience. As an example of needed modification we have, by permuting rows or columns once, algebraically written

$$I_1 Z_j(k) = \frac{1}{3} \left\| \begin{matrix} 1 & 2 \\ 2 & 1 \end{matrix} \right\|$$

a probability description of the 2-to-1 rule for favorites. Of course, the possibility of such manipulatively forced fitting does not validate the theory upon which the distribution is based. Interpretation of the algebraic operations used may give clues to a valid basis.

When $m = 2$,

$$Z = \frac{1}{10} \left\| \begin{matrix} 6 & 3 & 1 \\ 3 & 4 & 3 \\ 1 & 3 & 6 \end{matrix} \right\| .$$

This distribution and its successors $m = 3 \ldots$ have application in contest for rank as models for heuristic discussion of the case of selected and controlled contestants. Pearson's Tables give Z for larger values of m.[60] In general, the chance of rank $k = 0$ for the favorite is never less than even

$$Z_0 (0) = \frac{m+1}{2m+1} > \frac{1}{2} \qquad m = 0, 1, \ldots$$

For unselected aggregates of contestants and races, the corresponding relative frequency tends to be less than .5, around ⅓, as we have seen.

Although Laplace and later students of inverse probability had no recorded interest in orders of succession in contest for rank, or in the sport of kings, it is a consequence of the Laplace distribution that the probability that the j-th favorite will be led at the finish by presumptive followers, is given by his controversial "law of succession":

$$\frac{j+1}{m+2} = \frac{1}{m+2}, \quad \frac{2}{m+2}, \quad \ldots$$

$$j = 0, 1, \ldots$$

Inversely, the probability that the k-th leader at the finish was a non-favored entry is

$$\frac{k+1}{m+2} = \frac{1}{m+2}, \quad \frac{2}{m+2}, \quad \ldots$$

---

[60] Pearson, K., Tables for Statisticians and Biometricians, *London*, pp. 81-97, 1914.

Fry, T. C., A mathematical theory of rational inference, *Scripta Mathematica*, Vol. 2, pp. 205-221, 1934.

Jeffreys, H., Theory of Probability, *Oxford*, 1939.

For purposes not concerned explicitly with contest for rank or races, Karl Pearson deduced other forms of Z for unequal ranges of j and k, and on more general assumptions; namely, that the prior "existence probabilities" are non-uniform. One or more parameters expressing this non-uniformity then appear with m in the formula for Z. Numerical expansions of these are also provided in his Tables. The Laplace form is analytically simplest and is perhaps the most striking of all models having an average linear relation between prior and outcome orders, symmetric variation from the average and parametric simplicity. The two "laws of succession" specify the linear discrete "regression" functions of outcome and prior rank.

We have thus indicated the descriptive uses of a highly adequate and flexible form $\Sigma n_i I_i/N$, with the unit distribution $n_0 = N$, and the null, (the distribution $n_i =$ constant) as extremes, and any desired "intermediate" variety of empirical or rational distribution. Laplace's is apparently the simplest intermediate example. It is clear that the possibility of its rational interpretation and empirical applicability does not validate its "inverse probability" basis. We may say, however, that it is possible so to select contestants and so to control actual social contest for rank that any distribution of outcome, however its specification was contrived, can be reproduced by observed frequencies of contest. We thus have a flexible descriptive codification of the statistics of contest for rank. As data accumulate we shall find biosocial explanation of special regularities. Mathematical probability rationale can make a contribution here, but we must not meanwhile mistake formal or statistical for social regularities.

9. INTEGRATION; ITS DEGREES AND LEVELS

It may have been noticed that the term organization, as distinct from the concept, has not been used without caution in the preceding discussion of the contest for rank. This vague term is avoided by treating each contestant as a unit, fully integral from beginning to end of the completed race, also treated as a unit and relying upon subsidiary notions such as "unity." However, we may differentiate contestants according to their unity; which might be defined in the race as the degree of mastery of the horses by their riders; and the m-races as complete, or not. Thus, the contestants, if first conceived as similar, stable and successfully functioning units, must be admitted to vary

in this respect when we consider them analytically. That there is up to that point a tacit assumption of comparable and sufficient internal unity, as we shall call it, ***integration*** of the contestants distributively, is clear. The race in which m is not constant may be considered if we wish.

It would seem to be understood and readily agreed that if a dual group such as the interspecies horse-man represents one level of integration, a cooperative contest or an alliance of two or more horsemen would represent another, of the next higher level. Similarly, rider and horse may be considered as units having a level of integration one lower than the reference level, horse-man, and so on, from level to level. Within each of these the units can be distinguished according to their unity or integration distributively.

The concept of intervals or levels of integration as distinct from degree of integration at a level has recently been developed rather fully.[61] Numerous levels are identified and labeled, from the earth and its biosphere upward to the galaxies, and downward to the "ultimate" or currently unanalyzed "particles" of physics. The originally hesitant resolution of the atom into electrons and complementary particles does not end the analytical process. Sub-electrons at first hypothecated have not been found necessary to explain observation.[62]

We can, conventionally, even short of compelling observations to the contrary, limit consideration to any level of integration as if it were ultimate and provisionally treat two successive levels as if further interpolation were not necessary, provided our framework receives without disruption any needed interpolations or extensions. Thus the viruses have recently been interpolated empirically between the living cell and its constituents on one hand, and the "inorganic" crystal on the other. The iso-electron is a possible conceptual extension, without empirical counterpart at present. Our basic problem here seems to be the construction of a complex framework to hold the incomplete series of specimens that has come to our attention empirically. The concept of levels of integration with its relations of spatial inclusion, of

---

[61] Footnotes 4, 7, 8, 9.

[62] Stranathan, J. D., The "Particles" of Modern Physics, *Philadelphia*, 1942. [The concept of levels of integration, as distinguished from atomic energy levels is not interesting to physicists generally, but the use of dramatic social metaphor among them is so well established that one questions whether a text on particle physics could be written without it.]

smaller to larger units, and of evolutionary succession in time is such a framework. That it has not yet been represented more formally or mathematically testifies to the inherent difficulty, and our inhibitions in the face of known and growing complexity at all levels.

The construction of mathematical rationale at or within a partially known level cannot await "all the facts" or logical reduction to the concepts applicable at other levels. A level will be characterized empirically in part by qualitatively or quantitatively invariant features—some of them specified by constant quality or metrical parameters; i. e., constant for that level. The latter have often been statistical probability measures of aggregates. By means of parametric measures of probability we shall here attempt so to express the integration of the "units." We shall first examine the basic problem of measuring integration between the members of a dual group at any one level of integration. Since the horseman is a classical representative of dual master-servant, leader-follower and contest relations, we shall use it freely in making our discussion concrete. Subsequently we can extend the method to multiple groups and a wider range of concrete instances.

10. DIFFERENTIATION OF THE DUAL GROUP

Consider a dual group (m = constant = 1) during unit time. Time is conceived to be measured on any appropriate or convenient physical or biosocial scale, i. e., not necessarily on that defined by celestial and clock motion, which is not always convenient, if appropriate. During this period let the group stand in either of the order relations represented by I_0, I_1 or else in a middle, transitional, or tied state which is neither one nor the other. We represent this by a momentarily undefined matrix I_{01}. Let the (positive) fractions of unit time spent in these states be a_0, a_1, a_{01} respectively. Then

$$a_0 + a_1 + a_{01} = 1 \qquad 0 \leq a_i \leq 1 \qquad (a_i \text{ not all zero})$$

Following a convention of classical probability theory, we now define the weighted sum

$$Z_j(k) = a_0 I_0 + a_{01} I_{01} + a_1 I_1.$$

It may be termed the "first expectation," or time average of the state variable I_i.

If we pick an instant in the unit interval at random, the probability that the group is in the state i is a_i, (on the assumption that all instants

are equally likely to be chosen). The fractional time interval a_i is thus a measure of probability that the group is in state I_i. We interpret Z.

In particular, when $a_{01} = 0$, zero time is spent in the transitional state, Z becomes

$$Z = a_0 I_0 + a_1 I_1 = \left\| \begin{matrix} a_0 & a_1 \\ a_1 & a_0 \end{matrix} \right\|, \; a_0 + a_1 = 1.$$

This is the form already familiar in the relative frequency representation of dual contest without ties (n_0/N, n_1/N). To interpret $a_{01}I_{01}$, imagine the contestants to be tied throughout the entire unit of time ($a_0 = a_1 = 0$). If instantaneous tied states are distributed into the classes I_0, I_1 by correlating them with ideally perfect coin tosses or by any mechanism characterized by equal probability for each of the possible resolutions, $a_0 = a_1 = \frac{1}{2}$. Then

$$Z = \frac{1}{2} \left\| \begin{matrix} 1 & 1 \\ 1 & 1 \end{matrix} \right\|.$$

This is evidently only one among possible distributions of resolved ties in a tied contest between equals. It is, however, consistent with our conception of equal chances of winning or losing as ruling in a "fair" contest between "equals." We can now drop the middle term $a_{01}I_{01}$ with the understanding that $a_0 = a_1 = \frac{1}{2}$ can represent the ideal state of equality of the chance of winning during a period taken as unity. The form of Z in which the middle term is kept explicit will be useful in an abstract theory of social groups. We shall not need it further in this paper.

The relation of mastery between man and mount, master and slave or servant, is often described by the couplet of terms dominance-subordination. The master (0) is said to dominate (control or order) the servant (1) who is subordinate to the master (controlled by, takes orders from), according to some empirical test of dominance-subordination here left unspecified. That such a relation can "vary in degree" [63] from "complete subjection" through "equality" and "overturn" or "revolt," including possible disintegration of the unit under critical conditions seems to be tacitly agreed also, That ideal case in which the titular master of a dual group, denoted $j = 0$, is invariably in "his" place $k = 0$, and the servant in "his" place $k = 1$, during the life of the unit, can be represented by $a_0 = 1$, or equivalently $Z = I_0$.

[63] e. g., Waddington, footnote 1.

Where the master dominates the servant "incompletely," their state can be represented $\frac{1}{2} < a_0 < 1$. Exact equality of dominance can be represented $a_0 = a_1 = \frac{1}{2}$. If in some actual group, either a_0 or a_1 be fairly extreme in value and stably established, it is theoretically possible to infer titular rank from the observed frequency n_i of the type of status behavior i in a relatively few tests ($N = 1, 2, \ldots$). Near-equality of status would be inferred from nearly equal relative frequencies of status behavior, n_i/N. Finally, where servant mainly dominates master $0 \leq a_0 < \frac{1}{2}$, revolt or overturn being defined in terms of a_0, time, and the turning point ($\frac{1}{2}$).

An adequate representation of group integration must provide for fancifully extreme states of inequality; i. e., judged by the run of our experience, for it may be called upon to describe or define Irving's "headless" horseman, as well as the rider who "gives the horse his head," is "mastered by" or "thrown" by the horse. These symbolize critical states where integrity of the unit is unstable and apt to vanish. When the titular head of a group "loses control" of the unit it may change heads or cease altogether to exist as a unit. Thus when it is uncertain during any period whether man or mount is in control (tentatively represented $a_0 = \frac{1}{2}$), we except their dissociation momentarily. Accordingly, $a_0 = \frac{1}{2}$ is in some sense a rational origin of measurement for integration of the dual group.

The parameter a_0 may thus be proposed as a "measure" of integration in some sense to be more fully defined for the dual group. If this proposal be rejected, there is no independent parametric measure available in our representation of $m = 1$, for ($a_1 - 1 = a_0$). We must abandon the prospect of measuring integration by means of a_0 (or a_1) and step outside this representation. If the proposal be entertained, any linear or non-linear monotonic function of a_0 (or a_1) may provisionally be considered as a measure of integration. In the former instance it remains to choose the origin and unit of measure.

We take $2(a_0 - \frac{1}{2})$ as our first measure of integration of the dual group. This makes integration zero at $a_0 = \frac{1}{2}$, and ± 1 as a_1 or a_0 vanish, where Z becomes I_0 or I_1 respectively. The definition is readily extended to multiple groups, as we show later.

If a_0 and $1 - a_0$ be chosen as measures of the relative dominance of the members $j = 0$ and $j = 1$ of the group, the measure of differentiation of relative dominance is $2a_0 - 1$. By this choice, dual group integration and differentiation of relative dominance are here not only

linearly, but exactly equivalent measures of the group. Consideration of an accepted verbal definition of organization [64] at this point will show that we have now distinguished and metricized two of its subsidiary notions. We now proceed to the integrative level of a group of dual groups and its integration.

The tacit assumption about a distributively qualified aggregate of contestants, as of horsemen, can be expressed metrically in terms of their internal integration, i. e., as integrated dual groups. The weakest form of specification of their metrical qualification is: the $N(m+1)$ contestants distributively have integrations which are non-vanishing. When we do not yet (as here) have an empirical measure for $(2a_0 - 1)$, we may assume it to be in fact within such a range as assures "successful" functioning. The contestants of our racing examples were of an assured type, for they were so selected by hindsight that all succeeded in crossing the finish line as units, that being one of the conditions for selection $m = 7$ of the 100 races from the record of all United States $m = 1, 2 \ldots$ races during June 1942.

Several empirical interpretations of $a_0 = 2a_0 - 1 = +1$ will now be attempted in order to fix our ideas of the numerical level of "nearly complete" dominance and "nearly unit" group integration. The meaning of "complete" dominance becomes clearer if we contrast it with the fair contest between two equals $a_0 = a_1 = \frac{1}{2}$. If, at the extreme, winning the contest were a complete certainty for $j = 0$, then $a_0 = 1$ and $a_1 = 0$, then losing it is a complete certainty for $j = 1$. Such extreme contests lack sporting interest, or for that matter any social function (except possibly the limiting one of defrauding bettors) since they only verify what was certain at the start. In the United States horse-race, for example, the extreme is approached by a starting price as long as 100 to 1 against the least-favored contestant, but infrequently, constituting only 5 per cent of the races included in our examples. Such contestants may be said to be "practically" out-classed. Odds of 1000 to 1 are nearly unheard of in single races, but in joint contingency play winnings at 1000 to 1 are recorded (see Tables 6 and 7).

It may be conjectured that "competent" riders and "well-broken" horses are so-called when in an integrative relation closely approaching $a_0 = 1$, the condition of complete dominance. How close this approach

[64] In the *Dictionary of Sociology*, organization is defined: That process which differentiates one part from another in a functional sense and which at the same time creates an integrated complex of functional relationships within the whole.

is numerically depends upon social context and the standard of dominance used in it, 100 to 1 or 1000 to 1 being figures that might emerge from statistical rider-mount breakdown studies. In master-slave or human-domesticated-animal groups, the master may have ultimate control over the life of the subordinate and literally "ride to death." Loss of life or any necessary function of one member of the group ends the life of the group. Too close an approach to equality may threaten or destroy its integration.

A possible form of the restriction upon real dual groups that emerges is therefore that extreme differentiation-integration be unlikely. That is, in real functioning groups the distribution of integration is such that

$$0 < (2a_0 - 1)^2 < 1.$$

For further important examples of the numerical degree of integration in real groups, namely, those characteristic of "fair," "good" or "sporting" chances in practice, we can turn to other sorts of human contest experiences. The favorites of our horse-racing examples turned in one victory in three races and the least favored about 1 in 100. By contrast in American baseball the favored club of the league of eight usually wins two out of three games, and the least favored one in three. These degrees of differentiation are historically fairly regular, expected and so well known as to have become almsot conventional conditions of this sport. If the distribution of victory becomes too unbalanced, reorganization of the clubs and leagues can be expected. We turn briefly to some social functions of sporting contest and their effect on *a*.

Elizabeth, in establishing the Queen's Plate Races, may have had a shrewd interest in "improvement of the breed." Such improvement can be determined by contest and promoted by training and selective breeding. One function of the race can accordingly be expressed as promotion of the differentiation of dominance and group integration by contest and genetic selection measured in terms of a standard or average specimen of a general population of men and mounts. At the same time, along with skill in equitation and contest and its symbolic and actual value in determining and maintaining intra-group dominance and caste relations, one may be promoting group chances of success in future military action by providing cadres or nuclei of organization.[65]

---

[65] Galton, Francis, Inquiries into Human Faculty [Queens' Plate Races], *London*, 1883, p. 329.

Veblen, Thorstein, Theory of the Leisure Class, *New York*, 1934, pp. 142-5, 276ff. [On functions of fast horses and gambling in maintenance of class relationships.]

These objectives can also be expressed $a_0 \longrightarrow 1$, referring either to the internal integration of Elizabeth's horsemen or to their dominance over possible others.

In brief, $2a_0 - 1$ measures differentiation of relative dominance within a group at a level and also the internal integration of the same group, e. g., of the horseman. As here defined, internal group integration and differentiation of dominance increase or decrease together for they are measured by the same number Referring to the next higher level, e. g., the group of competing horsemen, a_0 can also measure the differentiation of relative dominance and integration to each other of its sub-groups, horsemen.

As a result of choice of origin ($\frac{1}{2}$), our integration measure ($2a_0 - 1$) has a socially significant algebraic sign. In geometry such a quantity is represented by a directed line segment and called a vector. When positive this vector points toward $j = 0$, the titular head; i. e., the mainly dominant member of the group on the convention $a_0 > \frac{1}{2}$. When $a_0 = \frac{1}{2}$, $2a_0 - 1 = 0$, it points to neither (or to both) of the possible heads. Such a critical group may be said to be "two-headed," or "headless." The concept of partial orientation of a group toward its possible heads has thus been given a vector representation.

How long such limiting groups, $(2a_0 - 1) = (0, \pm 1)$, can function successfully in fact, depends upon the context. It may be that biosocial groups generally can be shown, on certain postulates, to have that permanence that leads us to distinguish them empirically from mere aggregates if and only if $a_0 \neq 1, \frac{1}{2}$. We propose later to distinguish "mere" aggregates and totally integrated aggregates from social groups generally in terms of their integration, as here defined metrically and thence extended to multiple groups.

II. INTEGRATION AND DIFFERENTIATION OF THE MULTIPLE GROUP

A formal representation that arises logically out of modes of record already well known in the field of social animal group studies or human sports is the pair-wise extension to multiple groups of the dual form already described in detail.

An aggregate of $(m + 1)$ social animals (i. e., of any internal integration distributively) is observed in paired contact interaction, at least once for each possible combination of units two at a time. In the simplest case, each interaction is classified dichotomously. This simpli-

fication expresses the primitive state of our knowledge and concepts rather than the nature of the content. The two classes of interaction, mutually exclusive and complete, are ($m = 1$) denoted ($0 > 1$), or ($0 < 1$) and ($j > k$) or ($j < k$), $j, k = 0, 1 \ldots m$ generally, i. e., when $m = 1, 2 \ldots$. In a sports context ($0 > 1$) may signify that $j = 0$ wins the contest, whence ($0 < 1$) that $j = 1$ wins and $j = 0$ loses, by the rules of play. In social animal studies these relations give place to behavioristic tests of dominance-subordination, as, the incipient attack and retreat reactions of animals or the active pecking of birds.

The two elementary forms of the dual representation of dominance are

$$Z_{j>k} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left\| \begin{array}{cc} - & 1 \\ 0 & - \end{array} \right\| \end{array}, \qquad Z_{j<k} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \left\| \begin{array}{cc} - & 0 \\ 1 & - \end{array} \right\| \end{array}.$$

The record of the peck order relation between $(m + 1) = 13$ domestic fowl observed by Allee is as shown in Table 8.

TABLE 8

Status of each individual in a group of thirteen brown Leghorn pullets on the basis of their observed peck-order

(R. H. Masure and W. C. Allee, *Auk*, 51; pp. 306-325, 1934)

Read the table: j is pecked by k, $j < k$, or k pecks j. A, BG, etc.; Masure and Allee code.

j		k	0	1	2	3	4	5	6	7	8	9	10	11	12
0	A		..	1	1	1	1	1	1	1	1	1	1	1	1
1	BG		0	..	1	1	1	1	1	1	1	1	1	1	1
2	BB		0	0	..	1	1	1	1	1	1	1	1	1	1
3	M		0	0	0	..	0	1	1	1	1	1	1	1	1
4	Y		0	0	0	1	..	0	1	1	1	1	1	1	1
5	YY		0	0	0	0	1	..	1	1	1	1	1	1	1
6	BG_2		0	0	0	0	0	0	..	1	1	1	1	1	1
7	GR		0	0	0	0	0	0	0	..	1	1	1	1	1
8	R		0	0	0	0	0	0	0	0	..	1	1	1	1
9	GY		0	0	0	0	0	0	0	0	0	..	1	1	1
10	RY		0	0	0	0	0	0	0	0	0	0	..	1	1
11	RR		0	0	0	0	0	0	0	0	0	0	0	..	1
12	RW		0	0	0	0	0	0	0	0	0	0	0	0	..

Because of the triangular relation between *M, Y, YY* this order is not "perfectly" linear. In the case of perfect linearity all unit cells would appear on one, and all zero cells on the other side of the diagonal. Such a linear order, in which each member of the group pecks the one next in the social order without being pecked by that member or the subordinates of the latter is called a "peck-right" order.

Pigeon flocks exhibit greater equality in the distribution of their dominance-submission interactions. It is necessary in such cases to observe a statistically large number of contacts between each pair to establish the existence of a permanent average order, if indeed any exists, during the time of observation. From Table 9 the differentiation of relative dominance, measured by twice the proportion of dominant contacts less unity, approaches zero. In the ideal case of equality the dominant interactions would divide equally and differentiations pairwise would vanish exactly. In the linear case, by contrast, the measure is either + 1 or — 1.

As an example of intermediate dominance we exhibit a typical table of games won and lost between eight clubs of a professional baseball league (Table 10). The social unit, the "club" of players, and the unit contact interaction, "game," are evidently much more complex than in the animal examples (Tables 8, 9). The difference is not only in the biological species of the units composing the club but in the range of levels of integration that become explicit upon closer inspection (birds, flock), (players, clubs, leagues). Formal identity of the representation derives from the primitive pattern of our social thought and its symbolic vehicles, as much as from the fact that both contexts are social group behavior.

This form can also be applied to the recording of individual or group preference;[66] with modification to the voting or like-mindedness records of councils and electoral, legislative or judicial bodies.[67] As a format for representing paired-order relations it is not, of course, limited in application to biosocial systems. The representative matrix algebra was already known, and used by physicists, before it became popular with psychologists and sociologists around 1930.[68] It is in

[66] Thurstone, L. L., An experimental study of nationality preference, *Jour. Gen. Psych.*, 1, pp. 405-425, 1928.

[67] See section 14 this paper for the *L* matrix.

[68] Guttman, L., An approach for quantifying paired comparisons and rank order, *Ann. of Math. Stat.*, 17, pp. 144-163, 1946. [Most comprehensive treatment to date.]

the latter contexts that functional as distinguished from mechanical order relations come into consideration.

To define the integration of a multiple group it is desirable to proceed now by logical extension of the dual case. If a_0 is the proportion

TABLE 9

Dominance frequencies of paired contacts between the individuals of a group of five pigeons

(R. H. Masure and W. C. Allee, *Auk*, 56; pp. 264-273, 1939)
GW, etc.; their code

Read: j dominated k, $n_{j>k}$ times, j dominated others $n_{j>}$ times
Total contacts for a pair $= n_{j>k} + n_{k>j}$

ABSOLUTE FREQUENCY

		k					
$n_{j>k}$		0	1	2	3	4	$n_{j>}$
GW	0	..	7	9	13	12	41
RY	1	10	..	8	10	10	38
BR	2	10	7	..	10	3	30
BW	3	4	13	7	..	8	32
RW	4	6	8	4	6	..	24
	$n_{k>}$	30	35	28	39	33	165

RELATIVE FREQUENCY

j	0	1	2	3	4	Relative rank
0	(.50)	.37	.47	.77	.67	.580
1	.63	(.50)	.53	.43	.55	.520
2	.53	.47	(.50)	.59	.43	.518
3	.23	.57	.41	(.50)	.57	.451
4	.33	.45	.57	.43	(.50)	.422
	Average relative rank					.500

Diagonal terms postulated.

TABLE 10

Paired comparison record of a season's play within a ball league
(Read: Games won by j from k; $j > k$)

CLUB j*	k								j	STANDING			RELATIVE RANK*
	0	1	2	3	4	5	6	7		Won	Lost	Played	
0	..	13	15	15	14	16	16	17	0	106	48	154	.688
1	9	..	14	15	16	16	16	18	1	104	50	154	.675
2	7	8	..	13	15	13	12	17	2	85	67	152	.559
3	7	7	9	..	12	9	16	16	3	76	76	152	.500
4	8	6	7	9	..	11	12	13	4	66	81	147	.449
5	6	6	9	13	11	..	9	14	5	68	86	154	.442
6	6	6	8	5	7	13	..	14	6	59	89	148	.399
7	5	4	5	6	6	8	8	..	7	42	109	151	.278
All										606	606	1212	.500

Ties: 4 — 7 (2); 1 — 2, 2 — 5, 3 — 6, 6 — 4, 4 — 3

Forfeit: 2 — 6 (1)

* j assigned according to relative rank.

of the whole unit of time spent by the dual group in state I_0, then $(1 - a_0)$ is the like proportion for state I_1. The foregoing example of near equality suggests as the complete form of $Z_{j>k}$

$$Z_{j>k} = \begin{matrix} & 0 & 1 \\ 0 & \| \tfrac{1}{2} & a_0 \| \\ 1 & \| 1 - a_0 & \tfrac{1}{2} \| \end{matrix}$$

The diagonal terms are specified conventionally to represent an ideally conceived dominance relation of equals, in this case of "self-identical" contestants. Empirical situations can be conceived and realized in which the postulated terms are approximately verified, as when the first and second teams of a nearly homogeneous club of players engage in intra-club games and divide victories about equally.

In all, $m\ (m+1)/2$ distinct dual groups can be formed by paired combination, and each such pair can be characterized by an integration measure. It is thus seen that our extended concept of the measure of integration of a multiple group is multi-dimensioned, even though we

can and may wish to define it for some purposes as some one of an infinity of simple mathematical combinations of the $(m + 1)^2$ component measures of the form $2a - 1$.

We shall here, as in the restricted definition, subtract ½ from each element of $Z_{j>k}$ and multiply each remainder element by 2. Extending this process to all $(m + 1)^2$ elements of $Z_{j>k}$, the set of integration measures may be symbolized by the matrix extension of $2a_0 - 1$.

$$I = 2\,Z_{j>k} - S_0/2 \qquad m = 1, 2, \ldots$$

where S_0 is a constant matrix, all of whose elements are unity. This reduces, as it should, to the more elementary definition when $m = 1$. It is seen that I and $Z_{j>k}$ are linearly equivalent measures of the integration of the multiple group just as a_0 and $2a_0 - 1$ are, since, except for a change of unit and complex origin of measure, $Z_{j>k}$, the multiple group differentiation of dominance is a measure of multiple group integration, if I is a measure. Thus multiple group differentiation and integration are equivalent measures. More general transformations of Z suggest themselves.

The handiness of one-dimensional numerical measures, and the familiarity of these in contexts where measurement is already a long-established and accepted practice, encourage unwarranted assumptions about the desirable or necessary nature of acceptable social measurement. Thus while it may be conceded on principle that social measurement, as, of organization, is likely to be "complicated," or even complex in the present technical sense, the irrational hope and expectancy of the handy one-dimensional measure lingers. Complex measures like I are apt to be unwelcome, however necessary on principle at the outset. Practical economy will continue to demand specification in terms of moderate aggregates of numbers, if possible, on some rational concept of equivalence between large aggregates of raw social measures and small aggregates of mathematical functions thereof.

The determinative elements of the reduced system of measures may possibly be viewed as measures of a set of "contributory factors" underlying the observed variation to be accounted for. Members of this set of factors can then be ordered by magnitude for priority of treatment in isolation, or neglected if of small relevance to some practical viewpoint chosen by or forced upon the observer.

We are thus brought to a more general concept and definition of multiple group integration. The scalar multipliers can be replaced by

matrices signifying a transformation of the accidental coordinate system and units of measure to that of some simple or more inclusive point of view. For economy's sake, we habitually seek to condense extensive tables of observations to a very few empirical or rational constants. Thus the Laplace distribution might serve to replace m^2 independent observed numbers by one rational parameter, the integer m, and the rules by which the general element of class (j, k) is generated. Without entering upon the advantage of others among the possible complex infinity of integration measures, the observer's record is itself an empirical, unreduced measure of differentiation-integration of the multiple group. Reductions need not reproduce exactly all the numerical specifications comprising the data for which they are substitutes. We recall here the chance or otherwise significant departures from the 2-to-1 rule or hypothesis for favorites. Rational schemes have an interpretive and heuristic value beyond curve-fitting, the Laplace distribution or its consequence—the Law of Succession in contest for rank being a case in point.

On our definition, integration and differentiation increase together linearly, or, are identical except for unit and origin and measurement. Evidently no such simple restriction is obligatory at the outset. We shall probably need to define differentiation otherwise after empirical investigation. For example, quadratic or power series expansions connecting them about a complex reference point should be useful on general principle. We see that integration and differentiation must be permitted an infinite variety of joint relationship on principle. While experimentally determined biosocial restrictions may demand joint linear increase in some or many cases, the arbitrary restrictions here imposed on our model do not further demand that any, let alone all, real social groups obey it. It is one manner of speaking about formally conceived groups that may have counterpart in reality. If not, the remedy is clear.

12. ORGANIZATION MEASURE

If a well-known verbal definition of organization as process is so acceptable as to be included in a recent dictionary of sociology,[69] it is worth an attempt at translation to our quantitative terms. It was with the view to expressing group organization in the form of a combination

[69] Footnote 64.

of subsidiary metrical terms, themselves specifying differentiation and integration of biosocial function that these terms were first expressed quantitatively for the group.

We have sought to formulate the manner in which "self-sufficient" parts become "subject to" or are "dominated" by an "over-riding" whole, as the biochemist Waddington puts it. These social metaphors are especially striking in a work on biochemistry, "Organisers and Genes." [70] We shall therefore hope that the method of our formulations may give more than a metaphorial comfort to biochemistry and morphogenesis. Of special interest here is Waddington's contention that "organisation is the fundamental notion and is capable of quantitative variation, whereas organism is not." This postulate is to the effect that organization is an intensive measurable state of organism or an extensive measurable magnitude associated with the process of organization in time.

It is conceivable that biology or sociology could long ago have proposed measures of functional organization as such, perhaps analogous to entropy in its early, non-probabilistic, sense. The nearest approach is Angell's complex index of social integration of American cities.[71] This is built up from various enumeratively determined demographic-social statistics. Its basic units are discrete and extensive, as social actions or enumerable persons or groups (welfare funds, people, spouses, families, companies); and its component indices are population rates or, ideally, statistical probabilities. But no mathematical rationale corresponding to that of entropy in early thermodynamics or later statistical mechanics yet inspires these measurements of social integration. "Laws" of "the" trend of social organization (e. g., social entropy) have been proposed, but not in terms of measures of these "entities." [72] Angell's social integration measure is statistical, as entropy measure and theory clearly became early in their history. Angell's measure was built up conceptually and from observations within the aggregates or groups that it measures, and without benefit of a precedent mathematical theory of social aggregation. While Angell's index is not offered as such it could be considered an index of functional organization on some extant definition (e. g., *Dictionary of Sociology*) provided that integration and differentiation of function have the same interpretation in both.

---

[70] Footnote 1. See also Schrödinger's use of social metaphor in the description of gene function (Footnote 14).

[71] Footnote 41.

[72] If entities they are. Footnote 39.

If not, we are left at least with the suggestion that functional organization measure could be defined as a mathematical composite of integration-differentiation measures, whether these are empirically or rationally defined.

We are thus brought to the concept of social organization as a ***mathematical function*** (not to be confused with function in our first sense) of variables measuring internal differentiation, and integration of the whole. This is analogous to the concept of entropy as a mathematical function of physical variables: of work done, heat energy exchange and temperature, or of the state probability (disorder) of the system. One difference is that we have not much restricted the class of forms, F, of the mathematical dependence contemplated. To state briefly that organization, denoted F, "varies in degree," and that it depends mathematically upon the measure of differentiation, D, and of integration, I, of the organized system, we can accordingly write

$$F = F\ [D, I]$$

where F [ ] is a symbol for the as yet unspecified form of mathematical dependence upon the measures D and I, themselves complex.

When in particular the system is a social group and, as we assume, integration is linearly equivalent to differentiation, D or I may be eliminated from F. The simplest form of dependence is then the identity of F and I: organization is equal to integration (differentiation), whose (complex) measure is I.

We can go on to express I in terms of $Z_{j>k}$ or Z. Thus, F, I and Z may all be considered as organization measures. Waddington's postulate "organisation varies in degree" will no longer be interpreted to mean that F is a *one*-dimensional measure. In particular, where $m = 1$, I reduces to $(2a_0 - 1)$. If we should agree on a standard reduction to dual form of all multiple groups, the scalar a_0 could be interpreted as a general measure of organization. In this case it would be a one-dimensional probability measure.

The restrictions upon F [ ] should not, as they were here, be chosen for mathematical convenience, but primarily under control of empirically determined biosocial restrictions. It seems probable that we shall require a measure of organization of two collectively isolated social groups, given their organization measures distributively. If it should for example be conjectured, and validated empirically, that "the

organization of two interacting social groups is the sum of their respective organizations," then their integrations, as here defined, do not combine additively but multiplicatively. A complex logarithmic form of F is then indicated, as was a simple logarithm in the case of entropy. But such analogy argument should never be mistaken for more than a clue to definition and redefinition of the terms of the analogy. If a calculus of functional organization emerges it will not then be merely thermodynamics in disguise.

We have accordingly tried to develop the idea of functional organization measure independently, rather than by pursuing uncritically some analogy with entropy measure (hitherto metaphorized "disorganization"). We have found that it can be expressed in terms of measures of integration and/or differentiation, which in their turn are complex probability measures of group state, as in the case of entropy. But organization is not reducible to entropy. Both may be components of a wider logical unity. There is nothing in our discussion that demands that, much less specifies how in advance of experience, "the" universal or local trend of organization should shape up as a function of time. Such demands for restriction upon the mode of temporal change, F, D, I, Z, etc., must come from outside our formal representation, whose aim has been adequacy in representing the facts of organization and its changes, without prejudice as to the actual course of change.

Our chosen measures of organization (i. e., of integration-differentiation) are already applicable at or above the level of the lower social animals, for indeed our constructive illustrations were drawn from these levels. The integrative level at which social life "really" begins in the sequence of levels can be considered in these terms.

The hitherto verbalized distinction between an aggregate and a group is expressible quantitatively for the first time in terms of our integration or organization measure. $I = 0$ may be taken to define the "mere" aggregate and $I \neq 0$, all integrated or organized aggregates. The fact that in a structural-mechanical sense stable wholes have been found, from the ultimate physical level upward, has led to the free extension of the concept and language of organization which belonged to the biosocial level originally. Physics has even been called the study of smaller organisms. This is not taken to mean that its particles are organisms in the functional sense, for the concept of adaptive function does not enter, except poetically, until we arrive at the level of the paracrystal and viruses at least.[73]

[73] Footnote 8.

If a mere aggregate be defined by $I = 0$ and all organized groups by $I \neq 0$, the problem of further differentiating groups according to "degree" of organization can no longer be conceived as a one-dimensional scaling of the degree of their organization, but as their approximate allocation to a representative manifold of complex numbers. The restricted sense in which, when $m = 1$, integration can be scaled on the interval $(-1, +1)$ or in which Z is "intermediate" to I_1 and I_0 for this case can be extended to the case $m > 1$. When $2Z_{j>k} = S_0$ $I = 0$. All the elements of I are at their extremes of magnitude distributively when $Z_{j>k}$ has unit or zero value for each of the above-diagonal elements, i. e., in the linear case. When the elements of $Z_{j>k}$ lie between 0 and 1 we may also say I lies "between" the possible extremes of I, in this extended sense. Taking I itself as a measure of multiple group organization, its specification includes that of $(m+1)^2$ scalar integration numbers, of which at most $m(m+1)/2$ are independent. When $m = 1$ integration of the dual group is a scalar

$$-1 \leq (2a_0 - 1) \leq 1$$

When $m > 1$ we can write symbolically

$$I_a \leq I \leq I_b$$

where I_a and I_b are symbols for the types of I analogous to the scalar extremes -1 and $+1$ for the integration of a dual group. This suggests that all multiple groups might be contracted to the dual form in a manner that permits comparison on the scalar interval of the dual group.

13. ORGANIZATION AND TIME

Evolution of organization can be discussed in terms of various possible specific forms of mathematical time trend. These statistical laws of predestination would contain pivotal constant forms of I, and time as the determining variable. In studies of social animal group organization, there is evidence for the need of such specifications of a regular process of maturation or decay. The author has already suggested general methods for generating time sequences of $Z_j(k)$ distributions.[74] Since we have accumulated few additional data to control our selection from the wide range of possible forms, we shall not detail these suggestions further.

[74] Footnote 22.

There is no space here to describe some preliminary studies of available time series of social group statistics. The data concern human sports and high order legislative and judicial groups. It is suggested that interested students turn their attention to the relatively simple evolution of a baseball league integration during season play.[75] If this seems lacking in serious import the trend of agreement-dissent and voting blocs in the United States Supreme Court, or the United Nations will provide ample challenge.[76] In this case the problems of accessions, or replacements of members, of invariance and change in the organization of a formally constituted group must be considered.

14. REPRESENTATION OF UNANIMITY AND DISSENT IN VOTING GROUPS

Voting has been defined [77] as "a procedure whereby group decisions are made the basis of choices exercised by individual members of the group . . . one of the available means whereby a society may organize itself and make formal decisions regarding power relations . . . for securing political obedience with a minimum of sacrifice of the individual's freedom . . . [its] essence is freedom of choice. Unanimous consent has been tried and found wanting." Majority rule is useful in ending deadlocks.

In order to represent the basic voting records of formal groups or representative bodies we must broaden our representation beyond the forms $Z_j(k)$ and $Z_{j>k}$. Zuben [78] is apparently first to consider the form of statistical representation of agreement to which he gave the name "like-mindedness distribution." It has been used by Pritchett [79] in studying the recent record of dissent in the United States Supreme Court. The record of voting of the Court during 1930-1945 has a formal parallel to that of dissent and blocing in the United Nations Security Council. In both, a persistent pattern of association into right and left blocs can be exhibited. We shall not enter far upon actual issues, constituency and voting records here. Our purpose is to set forth a formal representation of the voting of the fundamental

[75] U. S. Baseball Annuals.

[76] This paper, section 12.

[77] *Encyclopaedia Britannica*, 1946 ed.

[78] Zubin, J., A technique for measuring like-mindedness, (L_4 matrix), *Journ. Abn. and Soc. Psych*, 33, pp. 508-516, 1938.

[79] Pritchett, G. H., Ten years of Supreme Court voting, 1931-1940, *Southwestern Social Science Quarterly*, 14, pp. 12-22, 1943.

group of three sub-groups or their individual representatives upon a two-sided question. The latter may be taken as some proposal and any one of its alternatives, denoted p, $\backsim p$. The model of all procedural questions is whether the members of a group shall agree to be bound by a vote which is less than unanimous; e. g., 2 of 3 possible votes when $m = 2$.

We denote the voters $j = 0, 1, 2$; a vote for p by 1, and against p, or for $\backsim p$, by 0. Each voter may vote in two ways, leaving aside restrictions imposed by bloc membership; hence a group of three may vote in $2 \times 2 \times 2 = 8$ distinguishable ways upon a two-sided question. Abstention is not considered. The 8 possible patterns of votes are easily set out for $m = 2$:

VOTER j	i — PATTERN OF VOTES							
	0	1	2	3	4	5	6	7
0	1	1	1	0	1	0	0	0
1	1	1	0	1	0	1	0	0
2	1	0	1	1	0	0	1	0
For p	3	2	2	2	1	1	1	0
Vs p	0	1	1	1	2	2	2	3

When $m = 8$ or 10 as in the United States Supreme Court or United Nations Security Council, there are $2^9 = 512$ and $2^{11} = 2048$ possible patterns, assuming freedom from blocing restrictions.

The elementary like-mindedness matrix L_i states whether or not any voter k agrees or disagrees with a specified voter j. Complete agreement may be denoted 1, and complete disagreement 0, thus suggesting a representation for partial agreement in a possible extension to the case of partial adherence or fractional voting by delegations. The present convention of array designation $(j, k) = 0, 1, 2$ is understood. When $j = k$, $L(j, k) = 1$ to signify that each voter "agrees with self." There are four distinguishable pairs of like-mindedness matrices representing the 8 patterns possible of voting:

$$L_i = \left\| \begin{array}{cccc} 0 & 1 & 2 & 3 \\ \left\| \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{matrix} \right\| \\ 4 & 5 & 6 & 7 \\ \left\| \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\| & \left\| \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} \right\| \end{array} \right\|$$

Patterns 0 and 7 represent unanimity, 1, 3, 4 and 6 minority-majority patterns. Inference of stable bloc association or of relative dominance among groups is made not on the basis of single instances of $I_{1,\ 3,\ 4,\ 6}$, but on the relative frequency of recurrence n_i/N in time series of votes on crucial kinds of issues.

In the case of the United States Court, decisions are reached on about 150 cases per annum. These can be tabulated in the form $n_i\ L_i$ and summed over various legal, political or socio-economic sub-classes of issues, or classes of agreement L_i; and divided by Σn_i to give the proportion of such cases in which a specified Justice agreed with any other. Moderate aggregates of cases are sufficient to show the presence of more or less stable blocs and even predict trends in the association of voting in the Court during 1931-1940. though there was a complete turnover of personnel by the middle forties. An example is drawn from Pritchett's continuing study of the court since the controversies of the thirties. (Table 11).

Dissociation or differentiation of the justices along a scale of some sort is apparent. If their reputations as conservatives or liberals are recalled it will be seen that the table has revealed two main blocs (conservatives right, liberals left) with Chief Justice Hughes about mid-way between them.

Similar studies of the late League and the United Nations Security Council have yet to be made. For the latter, the data are sparse, especially when procedural and substantive questions are distinguished. Maneuver and debate mark this distinction, for the unstable relation of the powers is affected in crucial cases.[80]

---

[80] *United Nations Security Council Journ.*, No. 42, July 1946. The Spanish Question. Polish representative: "I think the question of whether this is a matter of substance or procedure is a very complicated one . . . ." Russian Delegate: "If this resolution is voted upon as a whole . . . I will vote against its adoption."

TABLE 11

Agreements among U. S. Supreme Court Justices in non-unanimous decisions, 1931-1935 terms *
(In percentages)

(Only non-unanimous decisions in which both members of each pair participated are included in the calculations.)

	St.	C.	Br.	H.	R.	VanD.	Su.	Bu.	McR.
Stone	..	89	83	54	43	33	29	20	22
Cardozo	89	..	79	59	42	36	31	21	26
Brandeis	83	79	..	60	55	41	39	31	30
Hughes	54	59	60	..	77	77	71	60	62
Roberts	43	42	55	77	..	74	78	71	67
VanDevanter	33	36	41	77	74	..	93	82	79
Sutherland	29	31	39	71	78	93	..	87	78
Butler	20	21	31	60	71	82	87	..	75
McReynolds	22	26	30	62	67	79	78	75	..

* See reference footnote 79.

The first mathematical theory of social decision by balloting appeared in France about the time of the Revolution.[81] Even today contributions to the subject retain a hesitant and primitive quality. Most of them are part-time diversions of scientists (Laplace) or dilettantes (Condoreet). During the period between World Wars a theory of power politics and armaments races was advanced apologetically and even defensively by L. F. Richardson,[82] as if his colleagues were hostile or loath to bother with the matter at all. We may say, however, that these mathematical theories of group decision by balloting and bulleting are

[81] Condereet, Marquis de, Essai sur l'application de l'analyse à la probabilité des decisions. *Paris,* 1785 (Stechert Reprint, *New York,* 1931).

Laplace, P. S., Philosophical Essay on Probabilities, Truscott-Emery translation, *New York,* 1902 [Probability of judgment of tribunals, pp. 132-139.]

[82] Richardson, L. F., Generalized Foreign Politics, *British Jour. Psych.,* Monograph XXIII, *London,* 1939.

off to a fair start, even though the facts they must overtake are far out in front.[83]

[83] Rashevsky, N., Mathematical theory of human relations, *Phil. of Science*, 2, pp. 413-430, 1935; and various issues of *Psychometrika*, 1, 1936-1942. [Economic, governing and other relations between two or more social groups.]

Thurstone, L. L., The prediction of choice, *Psychometrika*, 10, pp. 237-253, 1945. [In opinion polls, elections, consumer and other preferences.]

Penrose, L. S., Elementary statistics of majority voting, *Jour. Roy. Stat. Soc.*, 109-1, pp. 53-60, 1946. [Statistical theory of power of single vote, bloc vote, hierarchic group voting; sizes of groups, groups of spokesmen, world assembly.]

Richardson, L. F., The number of nations on each side of a war, *Jour. Roy. Stat. Soc.*, 109-2, pp. 130-156, 1946. [Statistics of wars according to number of groups allied.]

PENETRANCE

ESTIMATED BY THE FREQUENCY OF UNILATERAL OCCURRENCES AND BY DISCORDANCE IN MONOZYGOTIC TWINS*

BY GABRIEL W. LASKER
WITH AN EDITORIAL NOTE BY CHARLES P. WINSOR

IN a study of a number of traits of oral anatomy in a series of 338 adult white Americans including nine pairs of identical twins, a number of theoretical considerations have arisen. This note is concerned with the significance of studies of discordance in one-egg twins and unilaterality in single individuals for an understanding of penetrance. Environmental factors supervene on the outward expression of most hereditary traits, and the relative importance of environment and heredity in the makeup of the individual is a basic problem. One approach to this problem is through twin studies and the examination of laterality.

TWINS

The frequency of concurrent appearance of a trait in both members of pairs of one-egg twins (a pair of whom always have identical genotypes) provides an estimate of the extent to which a given trait is hereditary. Let a trait, A, occur with frequency p_A; then it will be absent (O) with frequency $1 - p_A$. If the trait is distributed at random without respect to genetic or other factors common to identical twins, one would expect the number of twin pairs of each type to occur with the frequencies p_A^2, $2p_A(1 - p_A)$, and $(1 - p_A)^2$. The extent to which a trait occurs less frequently discordant than the middle

* From the Department of Anatomy, Wayne University College of Medicine, Detroit, Michigan, and the Department of Biostatistics (Paper No. 236), School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Md.

term and more frequently concordant than the other terms is thus an "index of concordance." [1]

Take, for example, *torus palatinus*, a bony growth along the mid-line of the hard palate. I have observed it to be present in 5 pairs of identical twins, absent in both of 3 pairs and present in only one individual of one pair. In addition, Less ('34) illustrates the cross section of the palates of four pairs of identical twins, one pair of whom have a mid-line ridge and three pairs of whom do not. Thus the total incidence of *torus palatinus* reported in identical twins is $\overline{AA} = 6$ (46.2 per cent), $\overline{AO} = 1$ (7.7 per cent) and $\overline{OO} = 6$ (46.2 per cent) in which $\overline{AA}$, $\overline{AO}$, and $\overline{OO}$ stand for the number of concordant positive, discordant, and concordant negative cases respectively. *Torus palatinus* occurs in 169 (53.3 per cent, standard error 2.8 per cent) of 317 individuals selected at random as compared with 13 (50 per cent) of the 26 twins. Applying the former as a more representative estimate, one would expect by chance alone 49.8 per cent of the pairs to be discordant. We have actually one discordant pair in thirteen. The chance of as aberrant a result as this is about $(1 + 13) / 2^{13} = .0017$, so that it seems reasonable to conclude that there are factors present—presumably chiefly genetic factors—tending to produce similarity between identical twins in this respect.

To make an accurate estimate of the extent to which environmental factors may interfere with the phenotypical expression of the genotype in monozygotic twins, it is necessary to postulate several conditions. In a given series let *a* stand for the genotype ordinarily predisposing to *A*, and *o* for the genotype ordinarily predisposing to *O*. As the twins are "identical," all pairs must be *aa* or *oo*. The pairs which are included in $\overline{AO}$ therefore are also genotypically *aa* or *oo*. It follows therefrom that one individual in each such pair is either phenotypically *A* and genotypically *o* or conversely the other individual who is phenotypically *O* is genotypically *a*.

If the genotype of a trait may thus be represented by the phenotype of the opposite in individuals in discordant pairs, it may be expected to happen occasionally in both members of concordant pairs of identical twins, *aa* being represented by *OO* or *oo* by *AA*.

---

[1] In a recent study Schinz ('45) sets up formulae for estimating the degree of penetrance from the ratio of discordance to positive concordance in one-egg twins.

Consider the first alternative first. If some *a*'s are represented by *O*'s but *o*'s are represented only by *O*'s, then all *AA* pairs are *aa*, all *AO* pairs, being genetically identical, must also be *aa*, and *OO* pairs can be either *aa* or *oo*.

How under these conditions will *aa* pairs be distributed between the three possible phenotypes *AA*, *AO* and *OO*? Let $\bar{a}\bar{a}$ and $\bar{o}\bar{o}$ denote the total number of individuals in the same series who have the genotypes *aa* and *oo* respectively under the first hypothesis. Let the probability that an *aa* will be represented by *AA*, *AO* and *OO* be denoted by p_1, p_2 and p_3 respectively. In other words, $p_1 = \overline{AA}/\bar{a}\bar{a}$ and $p_2 = \overline{AO}/\bar{a}\bar{a}$ and $p_3 = 1 - (p_1 + p_2)$. If the likelihood of an *a* being represented by an *O* or an *A* in each of a pair of twins is independent of the chance in the other twin (a conservative hypothesis, for factors in the common environment may tend to produce positive correlation) then the probability, p_2, of type *AO* (including *OA* which is not distinguished) will equal two times the product of the square roots of p_1 and p_3; i. e., $p_2 = 2\sqrt{p_1}\sqrt{p_3}$.

We thus have three equations and three unknowns:

$$p_1 + p_2 + p_3 = 1,$$

$$p_2^2 = 4p_1p_3,$$

$$\frac{p_1}{p_2} = \frac{\overline{AA}}{\overline{AO}}.$$

Substituting the first and third of these in the second, we get easily

$$p_2 = 4\frac{\overline{AA}}{\overline{AO}} \Big/ \left\{1 + 2\frac{\overline{AA}}{\overline{AO}}\right\}^2,$$

and thence

$$p_1 = 4\frac{\overline{AA}^2}{\overline{AO}^2} \Big/ \left\{1 + 2\frac{\overline{AA}}{\overline{AO}}\right\}^2,$$

$$p_3 = 1 \Big/ \left\{1 + 2\frac{\overline{AA}}{\overline{AO}}\right\}^2.$$

To get the total of $\bar{a}\bar{a}$ pairs, we note that

$$\bar{a}\bar{a} = \overline{AA} + \overline{AO} + p_3\bar{a}\bar{a},$$

and hence that

$$\bar{a}\bar{a} = \frac{\overline{AA} + \overline{AO}}{1 - p_a},$$

or, after substitution and simplification,

$$\bar{a}\bar{a} = (2\overline{AA} + \overline{AO})^2 / 4\overline{AA}.$$

This gives us a maximum estimate for $\bar{a}\bar{a}$, in that we have assumed that all *oo* genotypes are phenotypically *OO*.

By an exactly analogous process one may determine a maximum probability for *oo* (and hence a minimum for *aa*) by assuming that all $\overline{OO}$ and $\overline{AO}$ pairs are genotypically *oo* and solving for the probability that *oo* pairs are *AA*.

In the case of *torus palatinus* $\overline{AA}$, $\overline{AO}$ and $\overline{OO}$ equal 6, 1 and 6 respectively, and the minimum and maximum probabilities for *aa* type identical twin pairs are hence at least 45.6 ± 13.8 per cent and 54.4 ± 13.8 per cent respectively.[2]

There are a number of difficulties about this method of estimation. The factor of sex may be important. Thus *torus palatinus* is particularly frequent in females (61.9 per cent of 168 females compared with 44.7 per cent of 159 males). Such factors can, of course, be taken into account, but *torus palatinus* is here merely cited as an example and detailed discussion of the trait is reserved for another occasion. Other complications are the difficulty of identifying one-egg twins, difficulties in identification of the trait in question, and a multiplicity of factors that affect random individuals differently from the way in which they affect twin pairs.

LATERALITY

There is another application of the principles here elaborated. It has long been known that inherited anomalies of bilateral structures tend to occur bilaterally. Differences between right and left sides in such cases may be considered as caused by unilateral modification of the phenotype from the genotype, a process analogous to discordance in one-egg twins.

Take eruption of the mandibular third molars, for example. In the 18 individual twins, both sides are erupted in 13, and one side in 3 and

[2] For a discussion of the precision of these estimates see the editorial note following this paper.

neither side in 2. Between twins the corresponding frequencies for the cross comparison (right side of one twin against left side of the other) are 13, 3, and 2, and for the homolateral comparison it is also 13, 3, and 2 respectively.[3] In 286 white Americans over 18 years of age the bilateral comparison yields the figures 179, 33, and 74 respectively. But the twins are all between 18 and 29 years of age and for this age group the ratio is 138, 24, and 64 respectively and there is no significant difference between the ratio of discordant eruption between one-egg twins and unilateral eruption in single individuals. In both cases the frequency is only a small fraction of what it would be by chance alone ($2p_A\,[1 - p_A] = 35.6 \pm 4.8$ per cent in the case of the 84 individuals over 28 years of age when eruption is presumably complete, while unilateral eruption occurs in 6.0 ± 2.6 per cent, a difference of 29.6 ± 5.5 per cent). It may be worthwhile to use the discrepancies between the right and left sides to estimate a minimum frequency of the genotype's not being represented by the phenotype. Applying to this situation the formulae already derived, it appears that in single individuals a mandibular third molar has the genetic tendency for eruption in 74.6 ± 4.8 per cent to 94.2 ± 2.6 per cent of cases.

Some traits are so rare that adequate series of twins are difficult or impossible to examine. In the case of *torus mandibularis*, a bony protuberance of the lingual surface of the horizontal ramus of the mandible, no case was found in the present series of identical twins. In 312 other adults examined, however, there were 51 (16.3 per cent) bilateral cases and 17 (5.4 per cent) unilateral. It may therefore be predicted that from 13.5 ± 1.7 per cent to 22.2 ± 2.4 per cent have the genotype for *torus mandibularis*, and at least 2.0 per cent of cases with this genotype may show no trace of the trait.

Actually, these estimations are not altogether satisfactory, for there are a number of untested presumptions. One of these is an equal sex distribution. This condition is actually satisfactorily met in the case of *torus mandibularis*, for the ratios of absence, unilateral and bilateral occurrences are 118, 8, and 24 compared with 126, 9, and 27 for males and females respectively. But it is also presupposed that unilateral occurrence is not influenced by any factor tending to make it appear more frequently on one side than the other. This is not true for *torus*

[3] All 5 individuals with unerupted mandibular third molars were under 22 years of age, and eruption cannot be assumed to be complete.

mandibularis; there are 15 occurrences on the right and only 2 on the left. Furthermore, of the 51 bilateral instances 19 are listed as more developed on the right and only 8 as more developed on the left.

There are implications for other problems in inheritance in the method here suggested for estimating discrepancies in the frequency of phenotype from genotype. It frequently happens that in apparently dominant traits an occasional generation is skipped, and in both dominant and recessive inheritance the frequency in brothers, sisters, and other relatives of cases differs from the Mendelian expectancy. If the environmentally produced differences in twins or on the two sides of the same individuals are adequate to account for these discrepancies, it is obviously unwarranted to search for more complex genetic explanations. In the usual case of this type there is no evidence of a certain trait in a specific proportion of individuals where it might have been expected to occur on the basis of a simple Mendelian explanation. If the discrepancy does not significantly exceed the limits set by the method here described, it may be ascribable to environmental influences.

On the other hand, if the discrepancy does exceed these limits, it still cannot certainly be ascribed to inheritance, for there are serious difficulties in the application of percentages for the genotype derived from identical twins or opposite sides of the same individuals to the wider problems of unrelated or more distantly related individuals. In most characteristics the similar post-natal environment in all but a few pairs of twins and the common pre-natal environment will tend to minimize environmental factors. In the cases of laterality studies the same criticism will generally hold, but not, perhaps, in such traits as those related to habitual use of one hand in which cases the error will be in the other direction. These objections, however, do not prevent the application of the method here suggested from clarifying the mode of inheritance. This will be particularly true of relatively common traits in which non-specific post-natal environmental factors play a large part.

SUMMARY

Several factors in the genetics of traits affected by environment are discussed. Firstly, if a trait is discordant in one-egg twins significantly less frequently than twice the product of the frequencies of the two types of concordant twins, then common factors, probably usually hereditary, are partially responsible.

Furthermore, a tendency which is known to be expressed in only one individual in some identical twin pairs may be genetically present in a certain percentage of pairs which show no evidence of the trait, or absent in some pairs with the trait. A method for estimating this possibility is presented.

An even more significant application of these methods to genetics is seen in the study of unilaterality. With the exception of traits influenced by specific genes for lateral dominance—and if allowances are made for limitations inherent in the data—much information on the degree of hereditability of a trait can be gained from a series of unrelated individuals.

Material must be carefully considered in order to discount many variables which might interfere with an estimation of penetrance. Of the oral traits considered in this discussion *torus palatinus* shows markedly different frequencies in males and females, eruption of mandibular third molars is progressive with age—especially in young adulthood, and *torus mandibularis* is more frequent on the right side of the mouth than the left.

The author wishes to express his thanks to Mr. Whitfield Cobb of Winston-Salem, North Carolina, who made helpful suggestions concerning the formulation of the equations.

LITERATURE CITED

LESS, LOTTE. 1934. *Gegenbaur's Morph. Jahrb.*, 74: 105-134.

SCHINZ, HANS R. 1945. Konkordanz, Diskordanz und Penetranz bei eineiigen Zwillingen (Versuch einer elementaren Darstellung), *Arch. der Julius Klaus—Stiftung;* 20: 13-25.

EDITORIAL NOTE

The preceding paper contains an interesting problem in probability which is worth a little discussion. I restate it, using a different notation.

A bag contains white and black balls, in proportions p and $1 - p$. A ball is drawn and shown to two observers, each of whom records (independently) his judgment of its color. Black balls are always recorded correctly. Each observer has a probability a of recording a white ball correctly, and $1 - a$ of recording it as black.

After N such drawings (with replacement), we have records in the form:

W cases recorded WW;
X cases recorded WB or BW;
Y cases recorded BB.

We should now like (1) estimates of the probabilities p and a, and (2) some measure of the sampling variability of these estimates.

We let w, x, y be the probabilities that a particular ball is recorded WW, etc. As estimates of w, x, y we have W/N, etc. We also have

$$\left.\begin{aligned} w &= pa^2 \\ x &= 2pa\,(1-a) \\ y &= 1-w-x = 1-2pa+pa^2 \end{aligned}\right\} \qquad (1)$$

as the relations connecting p, a with w, x, y. We get, solving the equations (1)

$$\left.\begin{aligned} p &= (2w+x)^2/4w \\ a &= 2w/(2w+x). \end{aligned}\right\} \qquad (2)$$

We may now substitute the observed proportions, W/N and X/N, for w and x, to obtain our estimates.

In Fig. 1 are shown the contours of p and a for the possible ranges of w and x. It will be observed that there is a region in the w, x plane for which (2) makes $p > 1$; and, in fact, that p becomes infinite for $w = 0$.

(In practice, I do not suppose that we should use (2) for estimation if the estimate of p were greater than 1. What we ought to do is a question I do not propose to discuss here.)

We may observe that there is no real difficulty here, so far as obtaining a point estimate is concerned. The values of w and x determine a unique pair of values p and a; no question of combining discordant observations arises, and there is no way of testing the goodness of fit by internal evidence, because there are no deviations to be accounted for. We further remark that equations (2) are the solutions of the maximum likelihood equations of estimation.

We are still faced with the question, "How good are these estimates?", or better, "What sampling variation should we allow for these estimates?"

For large samples, we may make use of the fact that these are maximum likelihood estimates, and compute their sampling variances and covariance by the methods appropriate to maximum likelihood.

We write:

$$e^L = C(pa^2)^W[2pa(1-a)]^X(1-2pa+pa^2)^Y$$

where $W + X + Y = N$ gives the distribution of observations in the three classes. Then, as is well known,

$$\partial L/\partial p = 0, \quad \partial L/\partial a = 0$$

are the estimation equations for p and a.

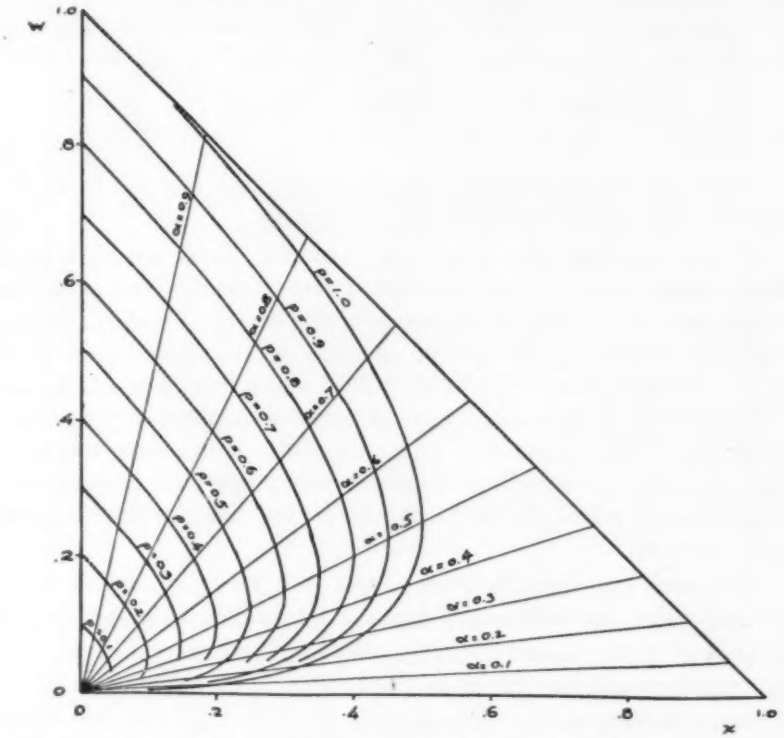


FIG. 1. CONTOURS OF p AND α FOR VARYING w AND x, FROM THE EQUATIONS

$$p = (2w + x)^2/4w,$$
$$\alpha = 2w/(2w + x)$$

Further, the information matrix is

$$I = \left\| \begin{matrix} -\dfrac{\partial^2 L}{\partial p^2}, & -\dfrac{\partial^2 L}{\partial a\,\partial p} \\ -\dfrac{\partial^2 L}{\partial a\,\partial p}, & -\dfrac{\partial^2 L}{\partial a^2} \end{matrix} \right\|,$$

with W, X, Y replaced by their expected values $Np\alpha^2$, $2Np\alpha\ (1-\alpha)$, $N(1-2p\alpha+p\alpha^2)$; and the variance-covariance matrix is obtained by inverting the information matrix. The details of the algebra need not be given; the results are as follows:

$$\left.\begin{aligned} \sigma_p^2 &= \frac{p}{N\alpha^2}(1-2\alpha+2\alpha^2-p\alpha^2), \\ \sigma_\alpha^2 &= \frac{(1-\alpha)\ (2-\alpha)}{2pN}, \\ \sigma_p\sigma_\alpha\rho_{p\alpha} &= \frac{-(1-\alpha)^2}{N\alpha}. \end{aligned}\right\} \tag{3}$$

These, of course, are large sample results, and it is not at all clear how far they may be useful for small samples.

I have not been able to see any promising direct approach to the small sample case. Some possibly useful indications of what may happen may be obtained by the examination of special cases. We may begin by considering the specific example on the second page of the paper. We have observed 6 WW, 1 WB, and 6 BB. Suppose we ask: What values of p and α might have given these numbers with reasonable plausibility? Or, to put it a little differently, if we were testing the hypothesis that p and α had numerical values specified in advance, in what range of values of p and α should we feel that the observed result was reasonable?

If p and α are specified, the χ^2 test is a very natural one to use. For any particular values of p and α it is not difficult to calculate the value of χ^2 for the sample (6, 1, 6). We can, then, determine the (p, α) contours for particular values of χ^2. In Fig. 2, the contour for $\chi^2 = 6$ is drawn, corresponding approximately to $P = .05$. For this sample, any values of p and α within the contour would give $\chi^2 < 6$, and would therefore be accepted as reasonable at the .05 level of significance.

It is of some interest to compare this contour with a crude application of large sample theory. From equations (2) we have estimates

$$p = \frac{13}{24}\ , \qquad \alpha = \frac{12}{13}.$$

Substituting these in the expressions (3), we have

$$\sigma_p^2 = \frac{67}{3456} \quad , \quad \sigma_p = .13924;$$

$$\sigma_\alpha^2 = \frac{168}{28561} \quad , \quad \sigma_\alpha = .07670;$$

$$\sigma_p \sigma_\alpha \rho_{p\alpha} = \frac{-1}{2028} \quad , \quad \rho_{p\alpha} = -.0462.$$

Now the correlation between p and α is low, and we shall neglect it. Then large sample theory says that the frequency surface will have elliptical contours with major and minor axes in the p and α directions,

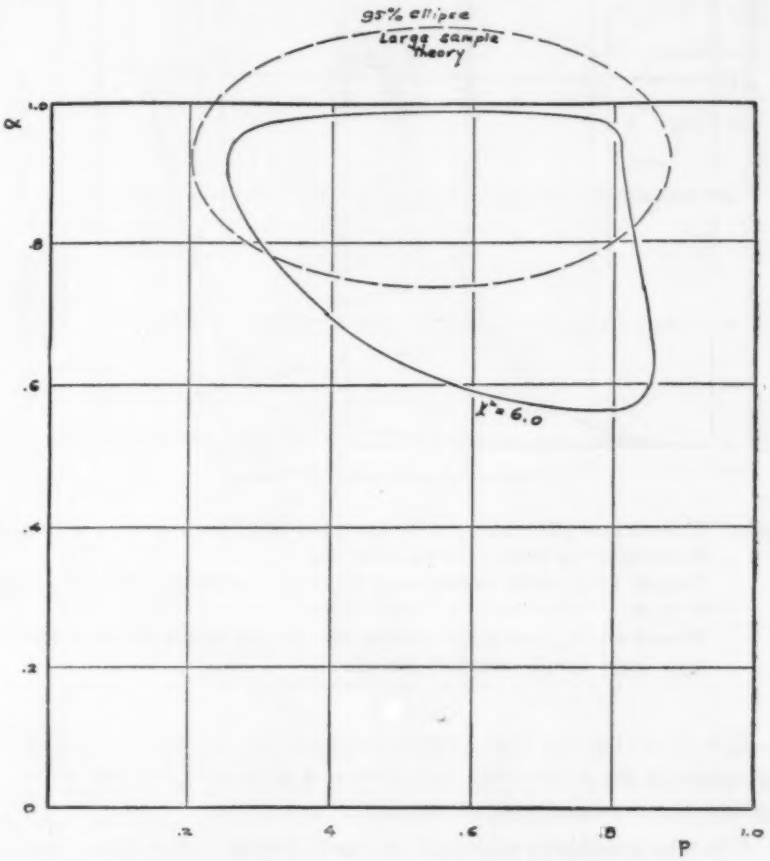


FIG. 2. "CONFIDENCE LIMITS" FOR p AND α, FOR A SAMPLE (6, 1, 6)
Solid curve: contour $\chi^2 = 6$.
Dashed curve: 95% ellipse from large sample theory.

centered at $p = \frac{13}{24}$, $a = \frac{12}{13}$. In particular, the ellipse enclosing 95 per cent of the total frequency will have semi-axes 2.447 times σ_p and σ_a. This ellipse may now be compared with the contour previously obtained.

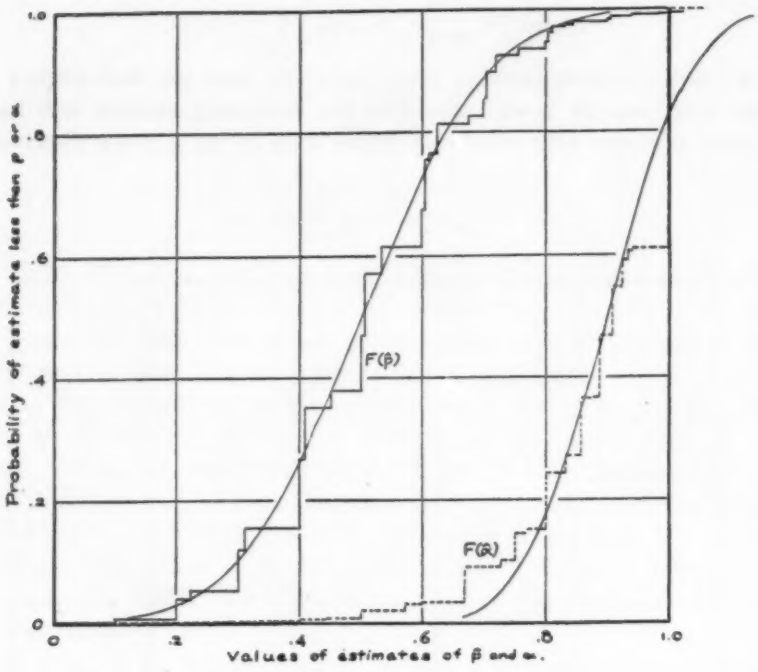


FIG. 3. CUMULATIVE PROBABILITY DISTRIBUTIONS OF ESTIMATES OF p AND a, FOR SAMPLES OF 10 FROM $p = 0.5$, $a = 0.9$
Stepped line: exact cumulatives; $F(p)$ = probability that the estimate is $\leq p$.
Smooth curves: cumulative normal curves, with means at true values, and with large sample standard deviations.

It is plain that the large sample theory is not too good; there is fair agreement in the p direction, but there is a large area for low values of a which lies outside the large sample ellipse.

It is also possible to approach the small sample problem by considering the actual probability distribution of estimates in sampling from a known population. For, knowing the true values of p and a, we

can calculate the probability of getting any possible sample, and hence the probability of each possible estimate for samples of any size. The computations offer no difficulties other than the mechanical labor involved. I have carried out the computations for two cases:

$$p = .5, \quad a = .9, \quad N = 10 \quad \text{(Fig. 3)},$$
$$p = .5, \quad a = .5, \quad N = 10 \quad \text{(Fig. 4)}.$$

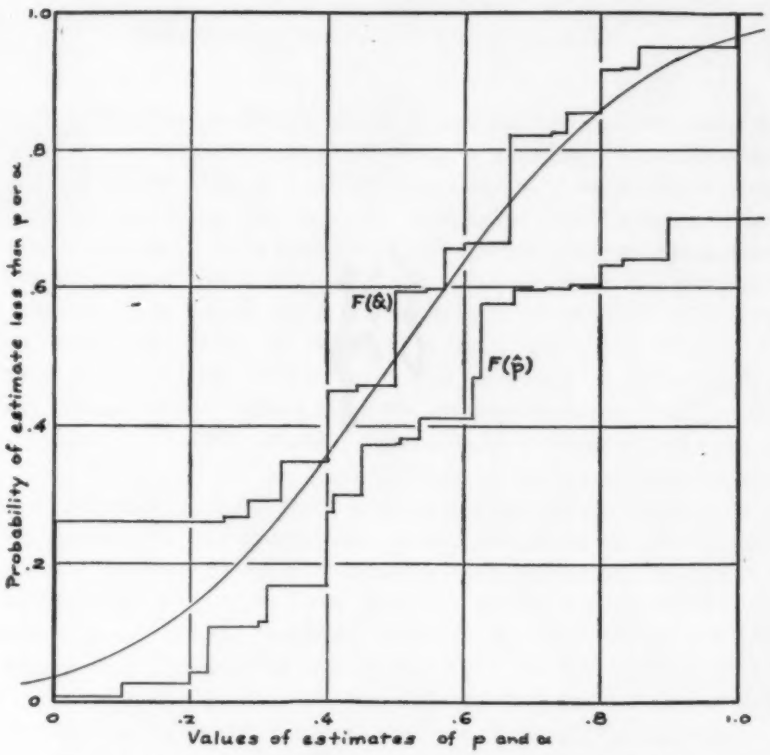


FIG. 4. AS FIG. 3, FOR SAMPLE OF 10 FROM $p = 0.5$, $a = 0.5$

Figs. 3 and 4 show, for these two cases, the cumulative probability distributions of the estimates of p and a obtained from equations (2). The stepped lines give the probability of obtaining an estimate less than or equal to any specified value; thus, for example, from Fig. 3, the probability that the estimate of p will be less than or equal to 0.5 is 0.47 (more accurately, 0.47096).

To permit a comparison with large sample theory, there are also shown cumulative normal curves (with standard deviation from the large sample formula) centered at the true values of p and α. It will be seen that in Fig. 3, where α is close to 1, the normal curve gives a good approximation for p, but a less good approximation for α. In Fig. 4, where $\alpha = 0.5$, neither distribution is even approximately fitted.

C. P. W.

EVALUATION OF SLOPES AND INTERCEPTS OF STRAIGHT LINES

BY LOIS M. ZUCKER

Department of Pathology, Columbia University, New York.

MANY scientifically useful quantities are estimated not by direct measurement, but must be calculated from the slope or intercept of a straight line fitted to a set of observations on two related variables. Fitting by eye is frequently unsatisfactory because of the multiplicity of possible lines and the absence of any estimate of precision for values calculated from the particular line chosen. The physicist, chemist, biologist or engineer consulting available reference works in search of a better solution is likely to be either seriously misled or else very much confused if he consults a variety of authorities. Because of the general confusion and because the problem is not fully solved, it seems to be in order to consider in one place the properties and inter-relations of the various procedures used or recommended and their possible adaptability to actual experimental conditions. An attempt will be made to make the presentation simple and intuitive rather than rigorous, and geometrical in approach as far as possible so as to focus attention on the application to what is essentially a graphical problem—curve fitting. We will assume that we are satisfied that the data actually represent (or that background information requires) a straight line and not a curve. Also we will assume that all values of Y are to carry equal weight in the solution, and the same for all values of X.[1]

The equation of the line to be fitted is $Y = kX + I$. (See Fig. 1.) An observed point X, Y will deviate from its correct position due to error by a line of length D. For a ready understanding of the relations among the estimates of slope or intercepts obtained from the different

[1] This last assumption is of course not true of many sets of data; however, when this more limited problem is solved, it can presumably be generalized without too great difficulty. The principles of the problem are illustrated just as well by the simpler situation where all points have equal weight.

solutions to be discussed, it is convenient to choose new variables x and y obtained by transferring the origin of the coordinate system to the point representing the mean of all X's and the mean of all Y's. All recommended solutions pass through this point, so that I has the value $\overline{Y} - k\overline{X}$, using the appropriate k for the particular solution chosen. (Many people fitting by eye do not go to the trouble of locating this point, although probably willing to admit that a good choice of line should pass through it.) In the new system the equation has the form $y = kx$.

Several of the lines to be discussed are based on the principle of least squares. If the errors which cause each point to deviate from its true position are random, independent, and each follows a normal Gaussian distribution law, then it will be true that the sum of squares of the combined errors in X — as measured from the true positions — will be less than if the errors are measured from any other positions, and similarly for errors in Y. (This will be exactly true with an infinite number of points, and true within sampling error for a finite number of points.) The solution follows from setting up an expression (one for each point) for the sum of squares of the deviations, properly measured and combined, and minimizing the value of this sum with respect to the value of the desired slope,

$$\frac{\partial}{\partial k} \sum_{x,y} (D^2) = 0.$$

We shall discuss four basic least squares lines, with special cases of some of them, and designate the estimates of D and k associated with these lines by appropriate subscripts: subscript y for deviations measured in the Y direction; subscript x for deviations measured in the X direction; p for deviations measured perpendicular to the line; 45 for deviations measured in the direction midway between the X and Y directions. These points are illustrated in Fig. 1. The symbol D will be reserved for the true deviation; the conditions under which D may be represented by D_x, etc., will be discussed later.

Table 1 summarizes these four basic least squares lines, showing in the second line the expression for D; the sum of the squares of D, minimized with respect to the value of k, leads to the expressions for k shown in the bottom line. In addition to these four lines, and variants on them, we shall consider solutions resulting from the procedure

called the method of averages. Table 1 also summarizes various recommendations as to usage from selected but, it is hoped, representative sources.

Characteristics of the lines obtained by minimizing the sum of squares of the deviations. They are a set of lines all passing through the mean of the observations. They are identical in slope only if

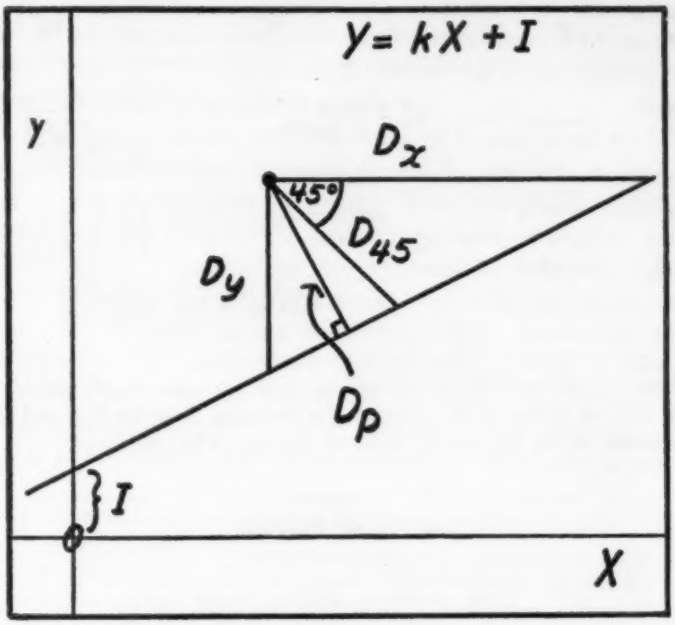


FIG. 1. DEFINITION OF TERMS

there is perfect adherence of all the points to a linear trend; for then, since $y = k_y x$ is exactly true for all x, y pairs, $\Sigma xy = k_y \Sigma x^2$ and $k_y \Sigma xy = \Sigma y^2$, so that $\Sigma xy/\Sigma x^2 = \Sigma y^2/\Sigma xy$, and it is seen from Table 1 that this means that $k_y = k_x$. With perfect fit, associated x's and y's always have the same sign so that there is no cancellation of terms due to opposite sign in adding up the xy terms, and Σxy has its maximum value of $\sqrt{\Sigma x^2 \Sigma y^2}$. With random scatter there is no association of sign whatever and Σxy approaches zero. When Σxy is zero, k_y must be zero, k_x infinite, and k_p and k_{45} indeterminate. With moderate scatter, since Σxy is always made smaller by cancellation of terms than it would be if there were no deviations from the linear trend, k_y, with

TABLE 1

Various regression lines and their characteristics

LEAST SQUARES LINES				
Direction of deviation	y	x	p	45° *
Single deviation (D)	$y - k_y x$	$x - y/k_x$	$y \cos \alpha - x \sin \alpha$ $\alpha = \tan^{-1} k_y$	$\frac{\sqrt{2}(y \cos \alpha - x \sin \alpha)}{\cos \alpha + \sin \alpha}$ $\alpha = \tan^{-1} k_{45}$
Slope (k)	$\Sigma xy / \Sigma x^2$	$\Sigma y^2 / \Sigma xy$	$\tan 2\alpha = \frac{2\Sigma xy}{\Sigma x^2 - \Sigma y^2}$ or $k_p^2 + k_p(1/k_y - k_x) - 1 = 0$	$\frac{\Sigma y^2 + \Sigma xy}{\Sigma x^2 + \Sigma xy}$, or $\frac{k_x + 1}{1/k_y + 1}$

* There are two possible 45° deviations. The one meant is the shorter one — i. e., the absolute value of the slope of the deviation direction is 1 and the sign is the opposite of the sign of the slope of the line to be calculated.

METHOD OF AVERAGES†
$k_A = (\bar{y}_1 - \bar{y}_2)/(\bar{x}_1 - \bar{x}_2)$

† The method of averages is older than least squares (1, p. 259). Some modern sources suggest it as convenient and simple as far as labor of computing goes, more satisfactory than fitting by eye, but inferior in correctness to a least squares solution. This sort of limited recommendation will not be further alluded to.

NOTES

Confidence limits are known for k_y and I_y, k_x and I_x (see, e.g., 5, 18, 13), $k_{p\text{-}\sigma}$ and $I_{p\text{-}\sigma}$ (15), and k_A and I_A (17).

Mellor (2), Wright and Hayford (3), Whittaker and Robinson (1) and Margenau and Murphy (4) recommend k_y as unique solution. This most usually used least squares slope may assume a more familiar look in terms of

this term in its numerator, must always be smaller than k_x with this term in its denominator. Thus the line with deviations measured in the Y direction is always flatter than the line with deviations measured in the X direction. The other two slopes k_p and k_{45} are always in between. The ratio $k_y/k_x = (\Sigma xy)^2/\Sigma x^2 \Sigma y^2$ goes from 0 (random scatter of points) to 1 (perfect fit to a straight line) and is therefore a measure of quality of fit to a straight line. It is a number whose value is independent of the particular units of the variables. The more usual measure of quality of fit in statistical discussions is the square root of this ratio, called r, the correlation coefficient.

Effect of improving the fit. The value of r, and hence the ratio of the two extreme least squares slopes k_y and k_x, depends on the eccentricity of the ellipse which can be roughly fitted to a graph of the points. If the eccentricity is increased, k_y and k_x approach each other, and this can be done either by extending the long axis — i. e., increasing the range covered by the observations — or by shortening the short axis, i. e., reducing the error leading to scatter. Both dimensions may be quite arbitrary as far as the basic straight line relation is concerned. The practical limitations on the extent to which the eccentricity can be increased, which always exist, may at any time be markedly reduced by changes in technique. If with a given error observations are made

---

the original variables if we remember that $\Sigma x^2 = \Sigma X^2 - (\Sigma X)^2/N$, $\Sigma xy = \Sigma XY - \Sigma X \Sigma Y/N$, etc.

Birge (5) recommends k_y not as unique, but because the various k's are not sufficiently different for typical physical data.

Ostwald-Luther (6) and Teissier (8) recommend k_y, k_x and their geometrical mean $\sqrt{k_y k_x}$, depending on the error distribution or on whether one variable can be regarded as directly dependent on the other.

Uhler (9), Camp (10), Kenney (13), and Wirth (14) recommend k_y, k_x and k_p, depending on the error distribution or the presence of an independently fixed variable. Greenwood and Yule (12) have a long discussion of the problem; they use k_p, admittedly as a reasonable approximation to the proper solution which they do not know. Yule and Kendall (11) refer to the case of errors in both x and y as an unsolved problem.

Worthing and Geffner (16) and Deming (15) recommend a special case of k_p, using special units for the variables; this, like k_p itself, reduces to k_y or k_x under appropriate conditions. It will be discussed below under the designation $k_{p\text{-}\sigma}$.

Ostwald-Luther (7) recommends $k_{p\text{-}\sigma}$, and also $\sqrt{k_y k_x}$, depending on what is known about the error distribution.

Wald (17) recommends k_A.

over a very short range, and then in steps the range is increased, the successive values of k_y and k_x, initially very different, converge steadily on some final value. The same process could be gone through by a hypothetical steady reduction in scatter from an initially very large one. It can be true for only one line (out of the set of all possibilities) that it shows no trend in value under such conditions, but merely fluctuates with steadily decreasing amplitude about the final result.

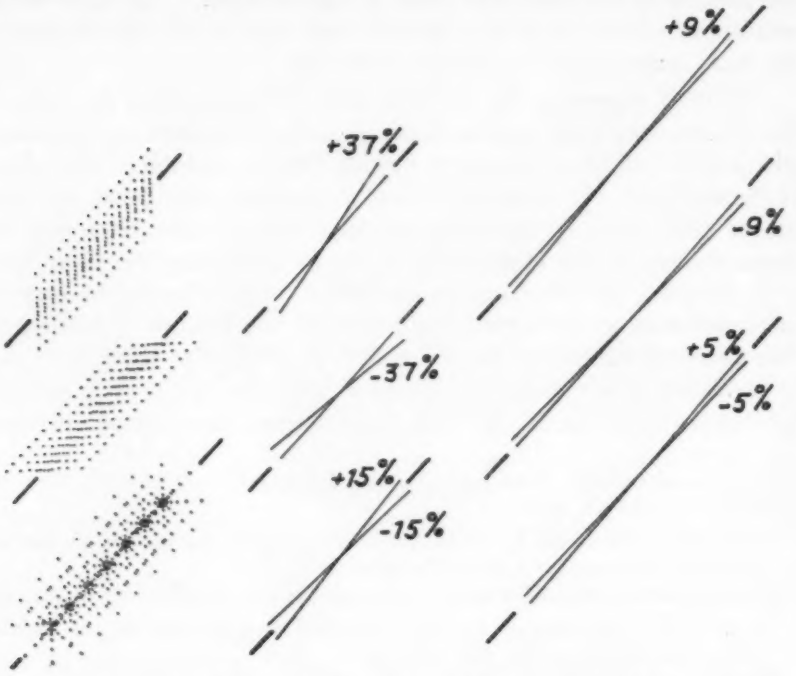


FIG. 2. SCHEMATIC SCATTER DIAGRAMS FOR VARIOUS ERROR DISTRIBUTIONS
Showing the corresponding pairs of regression lines (least squares lines with deviations assigned to the Y and X directions) compared with the true line; also the effect of increasing the range of variables covered on the difference between the lines.

This line is the solution we want to be able to calculate (representing for whatever degree of quality of fit an *unbiased* estimate of the desired parameter.) No other solution can fail to be biased, but the extent of bias will decrease steadily as the quality of fit improves.

The suggested procedure has been carried out for a few idealized distributions in Fig. 2. On top is a distribution which would result

from a number of observations of Y at each of a set of X values, where X can be accurately fixed, and errors in Y have a normal Gaussian distribution about their means. The means of the Y arrays coincide with the true values along the true line (indicated in heavy ink.) In the middle is the same situation for errors in X, and no error in Y. On the bottom is a distribution representing a system in which neither variable can be accurately predetermined, as for instance in observing two related properties X and Y of a set of biological organisms or chemical compounds or physical objects, the true points coming at regular intervals along the true line. With error in both X and Y, a set of observations on each individual would lie in approximately elliptical patterns around each true point. This reduces to a circular pattern if the X and Y errors are approximately equal, and to one of the two upper patterns if the error is almost entirely in one variable. To the right of the scatter diagrams are the two least squares lines with slopes k_y and k_x compared with the true line (heavy ink). Further to the right are the two least squares lines resulting from doubling the range, adding on in each case an identical set of points. For the three types of distribution, k_y, k_x and something in between are the desired lines.

With well-populated scatter diagrams as in Fig. 2, it is the shape at the ends which determines the behavior of the various lines. It can be shown that a least squares line calculated on all points individually is identical (within sampling error) with the same line calculated on properly weighted averages of the points, if the averaging is done in the direction to which the error is assigned in the particular least squares calculating being done. Averaging in the Y direction and in the X direction for each scatter diagram of Fig. 2 gives three pairs of curves (the regression curves of the statistician). Each pair is identical through the middle of the scatter diagram (and identical with the true line) where no end points are being averaged. Toward the ends for the top scatter diagram with vertical sides the averages in the Y direction will remain on the true line while the averages in the X direction curve up, on the right, and down on the left, thus swinging the line fitted to this curve around counter-clockwise from the true line. The end behavior of the regression curves is unchanged when the range is extended, but the central portion becomes increasingly longer, carries more and more weight, and eventually swamps out the end behavior. In the same way, in the scatter diagram with horizontal ends the

average in the X direction remains on the true line, and the average in the Y direction swings down, on the right, and up, on the left. In any scatter diagram with ends symmetrical with respect to the true line (when plotted with an approximate geometrical slope of 1), whether the ends have straight sides perpendicular to the line, or are curved or angular, both averaged curves deviate from the true line at the ends to about the same extent in opposite directions.

The distribution illustrated in the lowest graph of Fig. 2, in which X, Y observations are made on a set of different individuals, objects, or experimental setups, requires further comment. The case shown is one where the true points have uniform density along a segment of the true line, and are absent elsewhere. This is an idealized situation, which we can hardly expect to realize in practice (since the true values are unknown, and so not subject to selection). We may, and do, however, find actual distributions more or less approximating this. In particular, we are often in a position to control the general spread of the true values, and in considerable degree to control the relative numbers in different parts of the range. Thus, for instance, in data on the relation between two measurements on rats during growth (e. g., the data of Fig. 3c), if only rats of a certain age are considered, or only sexually mature rats, or only embryos, or only rats which have some definable and recognizable special genetic constitution among the many such making up the colony, there will be a different, smaller segment along the ideal line which can be represented than if all rats of all ages in the entire colony are included. Even the entire output of the colony represents a limitation of the same sort on the range obtainable if other genetic variations found in other colonies are included (supposing these do not upset this particular aspect of the rat body architecture which is being measured). Provided we have a class of individuals recognizable as in the same class by other criteria than the values of X and Y being measured, we will continue to have a situation basically like that illustrated in the lowest graph of Fig. 2, with the true points present on only a segment of the true line and the end behavior of observed points determined solely by the relative error in Y and X. If, however, the range is limited by including only those individuals lying between selected values of X or Y, the end behavior is of course affected by this choice; there will be either a vertical or horizontal cutoff, as in the upper two sections of Fig. 2, and the scatter diagram will look as

though in the end regions the error were suddenly all associated with Y or with X.[2]

Preliminary conclusions. In practice we frequently have only a poorly populated scatter diagram, and almost never a recognizable simple type. In practice we do not know the relative size of what we have called the end regions and the central region. In practice we do not know the true line or its orientation with respect to the scatter diagram, which we have assumed as known in this discussion of scatter diagrams. It is not being suggested that a proper choice of solution can be based simply on the appearance of a graph of the data. But if, from anything we can logically infer about the true points beyond the assumption that they lie on a line, and from the mode of collection of the data, and from any information about the nature and distribution of the error, we can guess what general type of distribution of points about the true line is to be expected, we will then at least have the problem set up rationally in that we will have a fairly good idea of whether k_y or k_x is a good estimate of the true value, or whether the true value lies somewhere between.

If there are so few points that the scatter diagram does not clearly belong to any particular type, and if it is not possible to collect more data, we shall have to rely on our general knowledge to select the proper solution. And if our general knowledge is inadequate for a decision, we shall have to recognize that any calculations we make will have not only the usual uncertainty which the statistician expresses in the standard error, but also the uncertainty arising from the differences between possible solutions other than the one we chose. More is said about this below.

It is evident that for the general case of one variable free of error the choice of solution can be affected by the mode of collection of the data, and simple situations result only when the end behavior is determined entirely by the error distribution (i. e., no selection), or by the compatible effects of the error distribution and the mode of select-

[2] Another possibility which should at least be mentioned is the situation which results from deliberate selection of values of one variable — say X — when there is actual error associated with that variable. In this case, if we select a particular value of X a number of times, the true values of X will vary, and there will be variations in the value of Y even if Y is measured without error. And, depending on the structure of the experiment, a number of interesting situations may arise.

ing the data. When there is error in both variables it is impossible for the end behavior imposed by fixing values of one variable to be compatible with the natural end behavior resulting from the error distribution. It is possible that under special circumstances, with error in both variables, k_y may be the best estimate of the slope. If there has been no selection of the points, the correct solution lies between k_y and k_x. We shall consider solutions for this case below.

Choice of line for prediction of one variable from the other. It is frequently stated in statistics texts that in order to predict values of Y for particular values of X the least squares line with slope k_y should be used, and to predict values of X for particular values of Y the line with slope k_x should be used, even though one or both of these lines differs considerably from the true line. This recommendation involves the use of two different lines for the same set of data, a practice which the physical scientist finds difficult to reconcile with his fundamental assumption of a fixed functional relation between the variables. It is true that the error of Y (given by the expression $\sqrt{\Sigma (D)_y^2/(N-1)}$) is less for the line with slope k_y than for any other line, and the error of predicting X is less for the line with slope k_x than for any other line, and this is all true within the same set of data.

But a glance at Fig. 2 shows how meaningless this slight superiority of prediction is if the range of true values and the variability covered are not fixed immutable properties of the relation under study; in the top section the prediction of X values for particular values of Y at the end of the range from the line with slope k_x is not on the true line and the prediction is not confirmed when the range is extended or variability reduced. The regression curve obtained by averaging all X's in each of several successive categories of Y gives an even smaller error of prediction than the straight line with slope k_x fitted to it. Now actually, in all but the end regions of the scatter diagram, the regression curve and the true line should tend to be identical while the regression line with slope k_x, being influenced throughout its course by the end behavior, has a steeper slope. For prediction purposes in all but the end regions of the scatter diagram, therefore, the true line gives a smaller error of prediction than the appropriate regression line, and as small an error of prediction as the appropriate regression curve. In the end regions the regression curve gives the best prediction, regression line next, true line worst. If the range of true values and the variability are fixed immutable properties of the relation under

study, then in the end regions prediction from the true line is biased, but if the range is not fixed (i. e., if we find that the definable group of individuals first studied is a sub-group in a larger one for which this relation holds) so that the X and Y values now at the end may be in the middle of the range with the next set of data, then the true line is better for prediction.

Error in both variables. Size and importance of the difference among the least squares lines. In a general way it is clear that if the scatter of the data is sufficiently small, the various least squares lines will be sufficiently alike so that it makes no difference which is used. Deming says: "If x has the same weight (i. e., the same precision) over all n points, and y likewise, x and y both subject to error, . . . the false solutions [our k_y and k_x] will hardly be distinguishable from the true solution [our $k_{p\text{-}\sigma}$] if the residuals are all fairly small." (15, pp. 185-186.) He gives no actual illustrations of what is meant by "small" residuals. Birge, writing for physicists, says: "I have found, with typical sets of data conforming to a linear equation, that the value of the constants given by [the two lines of slope k_y and k_x] differ by only about 1/5 of the probable errors of the constants. Such differences are of course devoid of any real significance" (5, footnote on p. 216). The particular data of interest in the discussion from which this quotation was taken are the Millikan oil drop data (19) which lead to the straight line plot shown in Fig. 3a. From the intercept shown the charge on the electron is calculated. The two extreme least squares lines are indistinguishable on this scale, and it can be calculated that the slopes and desired intercepts differ by slightly less than 1/5 of their probable errors. Not all data in the natural sciences are of the quality of these; Figs. 3-5 present other sets of data to which the least squares fit has been made or might be made. In all of these there is ambiguity in the choice of line to use, since X and Y are observed on a series of individuals and there is presumably error in both variables. In Fig. 4 the points represent carbon dioxide at various temperatures and pressures above the critical point, and X and Y are the fluidity and specific volume (20). In Fig. 5 the individuals are chemical compounds, X and Y are logarithms of rate or equilibrium constants of reactions (21, 22). In Fig. 3a the individuals are variably-sized variably-charged oil drops and X and Y are functions of their rates of fall under the influence of varying gravitational and electric fields. In Fig. 3b the individuals are again chemical compounds, and X and Y are functions

of their intrinsic viscosities and diffusion rates (23). In Fig. 3c and 3d the individuals are rats, and X and Y are the logarithms of their body weight and femur ash weight (24). In each case the two least squares lines obtained by minimizing the sum of the squares of the deviations in the Y and in the X direction have been drawn in. Since these are the two extremes of the group of possible least squares lines, the max-

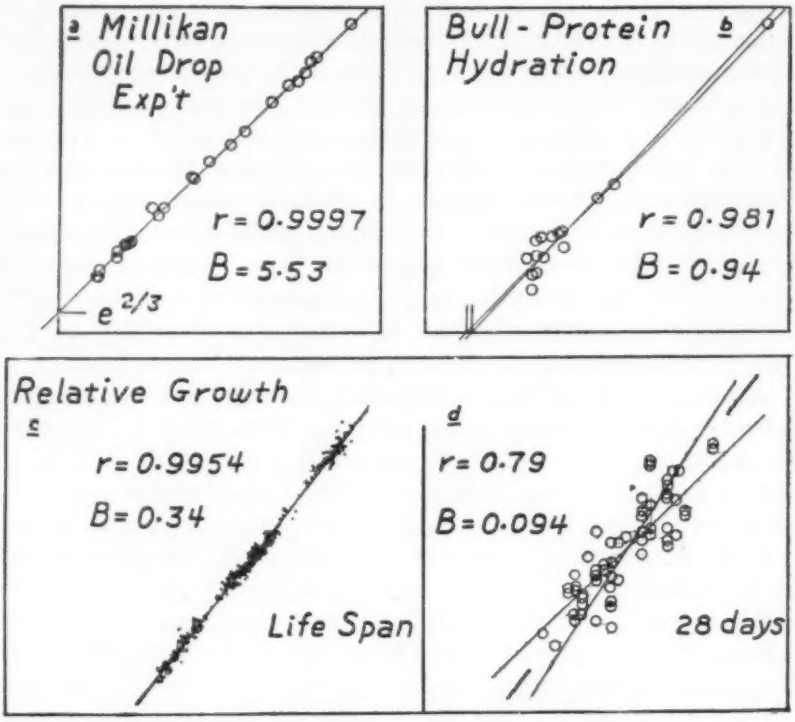


FIG. 3. REGRESSION LINES, y ON x AND x ON y

(a) Millikan's oil drop experiment (19), the charge on the electron being calculated from the 2/3 power of the intercept on the vertical axis.
(b) Data of Bull and Cooper (23) on the viscosities and diffusion rates of proteins in aqueous solution; from the left hand intercept on the horizontal axis the average degree of hydration of the proteins was calculated. The right hand intercept would give a somewhat different result.
(c) Log(femur ash) and log(body weight) of rats.
(d) Points from (c) for 28-day-old rats, on larger scale. Note the effect of range of variables covered on r and the difference between the two lines.

imal error which might be incurred would be the use of one of these two lines when the other is called for. It will be observed that this maximal uncertainty is of the order of the uncertainty associated with selecting the best fit by eye. In each case r is the square root of the ratio of the slopes of the two lines, the correlation coefficient.

The Birge criterion. In Figs. 3-5, B is the ratio of the probable error of one slope to the numerical difference between the two slopes,

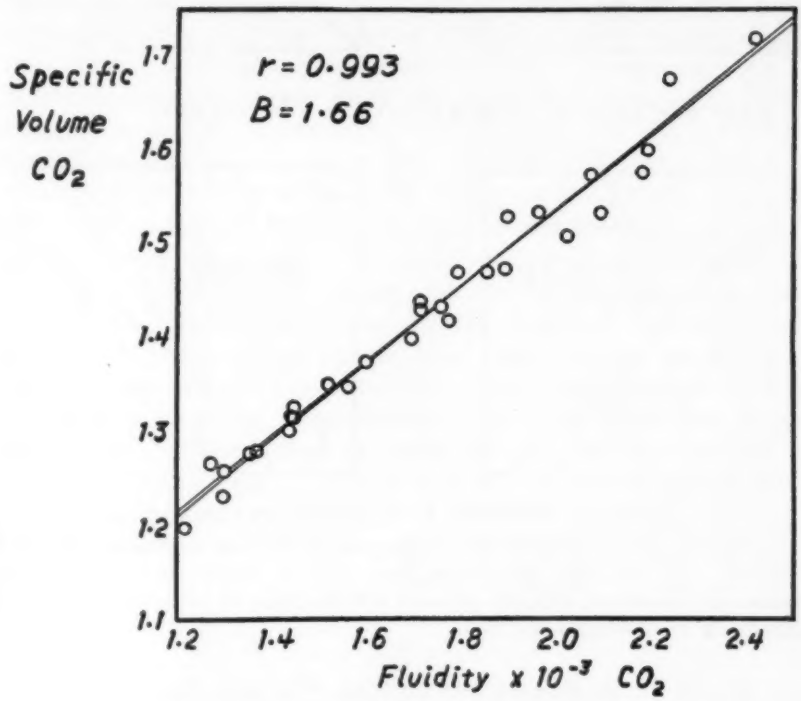


FIG. 4. REGRESSION LINES, y ON x AND x ON y

The relation between specific volume and fluidity of carbon dioxide at various temperatures and pressures above the critical point. The data are those of Phillips, as formulated by Jellinek (20).

a measure of the importance of the difference in slopes suggested by the remark of Birge quoted above. To quote Birge again, "One of the most fundamental rules of computing is that the error introduced by the calculation itself shall be negligibly small compared to the probable error" (18, footnote on pp. 4-5), and he argues that if the error introduced by the wrong choice of line is only 1/5 of the probable error,

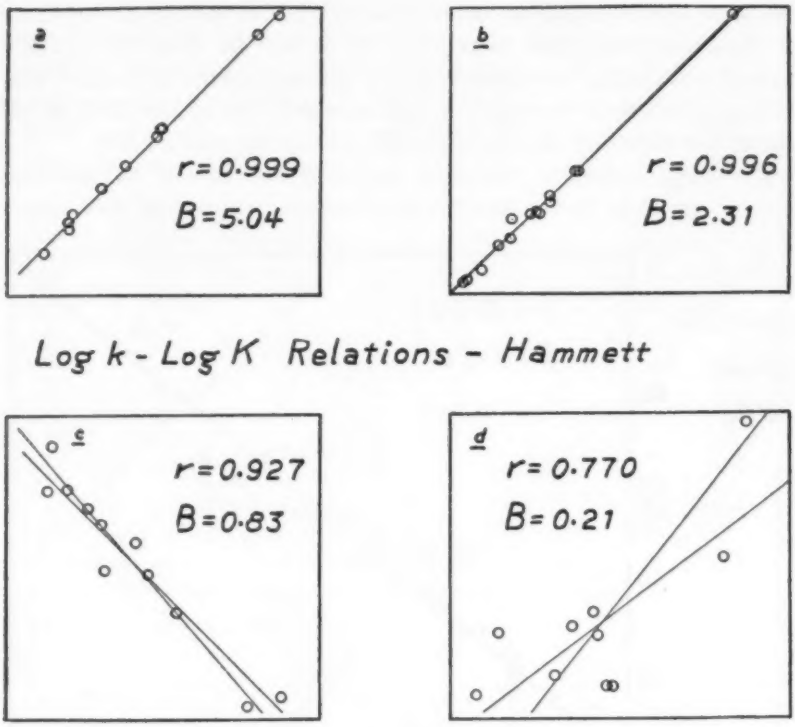


FIG. 5. REGRESSION LINES, y ON x AND x ON y

Double logarithmic relations among rates and equilibrium constants of chemical reactions. (a) represents the Kindler data (21) on alkaline hydrolysis of cinnamates and benzoates; (b), (c) and (d) are the reactions numbered 3, 18 and 8 respectively in Hammett's text (22).

as it was in the actual examples he tried, this error can be considered negligible. We can define

$$B = \frac{\text{P.E. of } k_y}{k_x - k_y} = \frac{.6745 k_y \sqrt{\dfrac{(1/r^2) - 1}{N - 2}}}{k_x - k_y}$$

$$= \frac{.6745}{\sqrt{\left((1/r^2) - 1\right)(N - 2)}}$$

where N is the number of observations. The formula for the probable error of k_y can be found in Kenney (13) or Birge (5). To meet the suggested criterion, B must be ≥ 5.

B varies in the expected way with quality of fit as measured by r; as r goes from 0 to 1, B goes from 0 to infinity, passing through the critical value 5 at some intermediate value of r fixed by the number of observations N. The nature of the dependence on N seems peculiar at first sight. The larger the number of points the smaller B is, and the harder it is for a given quality of fit to meet the criterion. Addition of more data points steadily, inevitably, reduces the probable error of the slope, even if the new points simply repeat the old ones. The probable error of the slope is just an indication of how reproducibly we can expect to repeat an experiment under the same conditions; the more points there are within an experiment, the more reproducible it presumably is, and the smaller the probable error. On the other hand, the difference between k_x and k_y is an expression of the particular distribution about the line exhibited by the points in a graph. This need not be affected at all by the addition of new points, and will not be affected to any extent unless the quality of fit is thereby improved. If there are relatively few points the precision will be poor, the precision measure so large that it will exceed the difference between the two slopes, and it makes no difference which line is used. As points are added the probable error goes down accordingly and eventually becomes too small to allow the Birge criterion to be met; where this happens depends on the quality of fit as measured by r. We can calculate the number of points which will allow a B of 5 for various values of the correlation coefficient, as in Table 2. Since to meet the criterion, B must be 5 or over, these N's represent the maximal number of points allowable for a given r if the criterion is to be met.

TABLE 2

Maximum number of observations permissible for $B \geq 5$, with varying correlation coefficients

r	0.9	0.95	0.995	0.999	0.9995	0.9999	1
N	2.2	3.0	4.0	11.1	24.7	93	∞

It appears from the values of B in Figs. 3-5 that the criterion is met ($B \geq 5$) among all these sets of data by Millikan's oil drop experiment and by the first of the set of log rate, log equilibrium constant data; therefore in many actual cases where one is interested in the precise

evaluation of quantities which appear as slopes or intercepts of straight line plots the choice of which line to use is of real practical importance.

The standard error of the slope. In reporting a result obtained from the slope or intercept of a straight line fitted to experimental data it is desirable to have some measure of precision. The uses of such a precision measure are the same as for precision measures of results obtained by simple averaging of experimental data, ranging all the way from a rough characterization of the quality of the figure reported to a measure which is needed to apply any of the statistical tests of significance. Significance tests are used to determine whether the figure being reported does or does not agree within its experimental error with another measure of the same thing obtained in a different experiment, or with some quantitative prediction based on theory. Whenever a least squares solution is used, and there is ambiguity about the direction in which the deviation should be measured, the precision measure for slope or intercept may be much too small and its use may lead to erroneous conclusions. In such cases the experimental uncertainty in slope or intercept value is not given by the probable error or standard error of the slope or intercept, but may be given by the full spread from k_x plus twice its standard error to k_y minus twice its standard error and similarly for the intercepts. For instance, Weymouth *et al.* (25), reporting on the relation between rate of oxygen consumption of isolated liver slices and the weight of the animals from which the tissue came, compare the results with a theory which calls for a straight line plot of certain functions of these observed quantities with a slope of $-1/3$. The least squares slope reported is $-.228$, with a standard error of $\pm.017$; applying Student's significance test they conclude that the observed difference of .105 in the slope could be expected by chance less than once in 10^8 trials, and that therefore the theory does not apply. From their results it is possible to determine that the slope reported is k_y (the Y axis being assigned to the oxygen consumption, the X axis to weight) and that the other slope k_x is $-.372$ with a standard error of $\pm.028$, also different from $-1/3$ but *in the other direction.* Some intermediate line would be identical with the theoretical prediction. There is no discussion in the paper to justify the selection of k_y; unless there is good evidence to support this choice the theory cannot be rejected as far as these data go.

The intermediate lines: dependence on choice of units. As can be seen from the expressions of Table 1, when the various slopes are

numerically much larger than 1, both k_p and k_{45} approach k_x in value, k_p more rapidly than k_{45}; when they are all very small in comparison with 1, both k_p and k_{45} approach k_y in value, k_p more rapidly than k_{45}. When the slopes are all of the order of 1, k_p and k_{45} approximate the geometrical mean $\sqrt{k_x k_y}$ in value, and the corresponding lines bisect the angle formed by the two extreme lines with slopes k_y and k_x. (These statements are exactly true when $k_p = k_{45} = 1$.)

Since the numerical value of the slope can be changed from very large to very small by suitable change in the size of the units in which the variables are expressed, and since the choice of units is usually a conventional one, k_p (and k_{45}) are unsatisfactory as such.[8] This behavior of k_p is readily visualized geometrically (the corresponding analysis for k_{45} is more complicated). If the units are such that the slope is very great, and hence a direct plot in these units is very steep, the direction perpendicular to the line is practically identical with the horizontal direction; if the units are such that a direct plot gives a very flat linear trend, the direction perpendicular to the line is practically identical with the vertical direction.

Deming's solution. Deming (15) presents a complete solution applying the least squares principle to the general problem of curve fitting which for the simple special case we are discussing (the curve to be fitted being a straight line, all Y values having the same weight, and all X values having the same weight), reduces to the line with slope k_p for a special pair of x and y units. The x and y units are to be in the ratio of the standard deviations (or probable errors) of X and Y, obtained under the same conditions of sources of error which apply to the data for the straight line relation being fitted. That is, in a series of N observations on X which should be identical except for the operations of all the various sources of error present when the data for the line are being collected, the observations will cluster around a mean which we can call $(\overline{X})_y$, signifying the mean of the observations when Y,

[8] Sometimes the relative size of the scale units is not simply conventional. This is true in the problem discussed by Greenwood and Yule (12), and in cases where Y and X are both logs. However, it is still true that if the error is largely with Y, the deviations should be taken largely in the Y direction, and vice versa. Therefore the only rational choice of unit for purposes of fitting a line is still based on the error distribution, whatever other rational choice may be suggested for expressing the underlying relation being studied. And if the different points carry different assigned errors, scale units must be separately adjusted — i. e., weighted.

whose variation also produces changes in X, is being maintained constant. Approximately 2/3 of all the X's will lie between $(\overline{X})_y - (\sigma_x)_y$ and $(\overline{X})_y + (\sigma_x)_y$, where $(\sigma_x)_y$, the standard deviation, is defined as $[\sqrt{\Sigma(X - (\overline{X})_y)^2/(N - 1)}]_y$. When the units chosen have the values $(\sigma_x)_y$ and $(\sigma_y)_x$, a given displacement in the direction of the one variable becomes of the same importance in indicating a real change as the same distance in the direction of the other variable.

The operation of expressing y in units of $(\sigma_y)_x$, x in units of $(\sigma_x)_y$, determining k_p for these units and converting k_p back to the original observational units is performed by using the equation

$$k_{p\text{-}\sigma}{}^2 + k_{p\text{-}\sigma}\left[\frac{(\sigma_y)_x{}^2}{(\sigma_x)_y{}^2(k_y)} - k_x\right] - \frac{(\sigma_y)_x{}^2}{(\sigma_x)_y{}^2} = 0$$

The proper root to be taken will be obvious for any particular problem. It will be found that substitution of dimensions for the various symbols in the equation for $k_{p\text{-}\sigma}$ (Y/X for k_y and k_x, Y for $(\sigma_y)_x$ and X for $(\sigma_x)_y$) leads to the proper dimensions (Y/X) for a slope, while this is not true for k_p as can be seen in Table 1. Comparing different problems with different relative errors in Y and X, it is clear that if $(\sigma_y)_x$ is very much larger than $(\sigma_x)_y$, $k_{p\text{-}\sigma}$ approaches k_y; that is, the solution approaches the one for error limited to the Y direction, as it should. And similarly if $(\sigma_x)_y >> (\sigma_y)_x$, $k_{p\text{-}\sigma}$ approaches k_x.

This solution is very attractive — it seems right intuitively. It is, in fact, so attractive, and Deming describes it in so pleasing a fashion that a careless student may expect more of it than it can actually give, or than any solution could give. This is the more likely to be the case when he finds Deming saying that the method not only finds the best values for the constants, but also finds estimates for the positions of the true points. In fact he says (15, p. 140), "Least squares is primarily a method of adjusting the observations, and the parameters enter only incidentally."

Actually, the "adjustment" does not give, and should not be expected to give, estimates of the true points in the same sense that the adjusted points are estimates of the true points when the error is all in one variable. That this must be the case is easy to see. We need only suppose that we are given the true line, and an observation lying on the line. No "adjustment" of this point is easily imaginable; and yet the true point may lie elsewhere on the true line. We cannot, it

is clear, expect to adjust for errors which displace a point parallel to the true line.

It may be helpful to look at the matter a little further. Suppose we transform the x and y scales to units of $(\sigma_x)_y$ and $(\sigma_y)_x$. In this scale, the distribution of errors around any true point will have circular symmetry. If we change to new coordinates along and perpendicular to the line, the circular symmetry will be maintained; and from a set of observed points we can get the errors perpendicular to the line, but we cannot even guess at the parallel errors. If we have only the observed points, and not the true line, we shall have to estimate the true line, which we shall not do exactly right, and we can now estimate the perpendicular deviations from the fitted line; but again we can do nothing about the errors parallel to the fitted line.

It is not to be supposed that what has been said is intended as criticism of Deming's solution, or of what Deming has said about it. It is intended to make clear what the solution does not and cannot do.

Difficulty in determining the relative size of the X and Y errors. Having come this long way, and arrived at a solution which seems to be right, it is discouraging to have to record that the solution is frequently unusable. The large practical difficulty with the Deming solution is how to estimate the quantity $(\sigma_y)_x/(\sigma_x)_y$. It is commonly the nature of the situation with both variables dependent that one cannot fix one variable in order to determine the error in the other. The $(\sigma_y)_x$ and $(\sigma_x)_y$ which can be determined as measurement errors usually represent only a small fraction of the sources of error leading to deviations from the line. To take the example of the log rate, log equilibrium constant relations of Fig 5, the immediate experimental errors of measurement — volumetric titrations, EMF measurements, etc. — can be estimated by taking a series of replicated observations. The effects of what might be called the secondary experimental errors of measurement — impurities in the reagents, competing reaction paths, side reactions, the magnitude of junction potentials, special salt or solvent effects, calibration or standardization errors, etc. — which in any particular determination will act as constant and therefore undetected errors, are extremely difficult to estimate; it is hardly possible to set up a series of observations in which these will all operate in random fashion.

A third type of error is the failure of the individuals concerned to meet exactly all the specifications for the particular linear relation

being fitted. In all of the examples of Figs. 3-5 there is essentially a set of individuals — oil drops, proteins, meta or para substituents in the benzene ring, containers of carbon dioxide, rats — with observations on two related properties of these individuals. It may be illuminating to regard all such individuals as being defined by a set of specifications, and as these specifications take particular numerical values or constellations of values one moves along the true line (26). In Fig. 4 there are apparently only two specifications, temperature and pressure, while the other cases are all more complicated. If there is no measurement error and if an individual truly fits the specifications, the result is a point exactly on the true line. Deviations from the line may be due to direct experimental error or to failure of the individual being observed to meet the specifications in all details. And the individual may fail to meet specifications either because it is not precisely what we think it is — due to chemical impurity or some other of the secondary type of experimental error — or because this individual in its pure form does not exactly obey the postulated relation. The points on the line may all be ideal points for which we can find only approximate actual examples. Thus the relations in Fig. 5 are regarded as resulting from direct linear relations between the free energy of activation and the free energy of the reaction as the structure of the molecule is changed to a strictly limited degree. For any substituent this change in structure means definite changes in energy distribution within the molecules of reactant, activated state and product; as long as these are of the proper kind the two free energies change concomitantly along the line. Hammett (22, pp. 121-124, 192) has pointed out that some of the deviations must probably be ascribed to second order changes in energy distribution which are not of the right type, of the sort which certainly becomes very large if ortho substituents are introduced, or if one moves the reacting group or substituent into a saturated side chain.

In the case of the biological data of Fig. 3, age and "ratness" are not the only specifications. For strict adherence some genetic differences are forbidden (since another colony of rats is known to adhere to a different line) although some seem to be allowable (since animals of all sizes within a colony adhere to the line, and the colony contains a considerable size range). To a certain extent it is possible to move along the line by varying the inherent size of the animals as well as by varying age. Experiment shows that certain even very mild nutritional

deficiencies are not allowable while others, even relatively severe ones, will affect the bone and body weight alike and are allowable. The rat must not have serious respiratory infections because such rats are consistently off the line. But among apparently normal, well nourished, healthy rats of similar genetic background there may be second order deviations from the proper specifications, just as among the class of meta and para substituted benzene derivatives there may be second order deviations from the proper specifications. How can we estimate the magnitude of these effects?

At any rate, this third type of error is absolutely tied up with the particular linear relation being considered; the error contribution of this type for a given measurement Y may be very different when it is paired with the associated measurement X', in a $Y - X'$ relation, than when it is paired with another associated measurement X'' in a $Y - X''$ relation. It must be concluded that we frequently do not have enough information to take advantage of the fine distinctions provided by Deming's solution.

The geometrical mean $\sqrt{k_y k_x}$. This is the solution recommended by the Ostwald-Luther handbook, 4th edition, for the intermediate case with error in both variables, and by the 5th edition for the intermediate case if nothing is known about $(\sigma_y)_x/(\sigma_x)_y$, so that $k_{y\text{-}\sigma}$ cannot be used. It can be derived as a special case from the general solution for measuring the deviations perpendicular to the line, if instead of choosing as units the error of observation of single values of X and Y, we choose as units the sigmas for the full ranges of the variables X and Y. The new unit is the quantity $\sigma_x = \sqrt{\Sigma(X-\overline{X}^2)/(N-1)}$, with Y allowed to vary over the full range, rather than $(\sigma_x)_y = [\sqrt{\Sigma(X-\overline{X})^2/(N-1)}]_y$ with Y fixed. In the terminology of Table 1, $\sigma_x = \sqrt{\Sigma x^2/(N-1)}$. These σ's will now include not only the deviations in X and Y due to error, but also the much larger deviations in each due to the concomitant variation of the other. Since they are measures of the range of X and Y covered by the observations, a plot in such units has the range covered by X occupying the same distance in the X direction as the range covered by Y occupies in the Y direction. Consequently the graphical slope of a plot in such units is 1. Therefore the geometrical mean slope is the one which would be obtained by converting to units representing the ranges of X and Y (which units will lead to a plot with slope approximately 1), determining the least squares line min-

imizing the perpendicular deviations from the line in these new units, and then reconverting to the old units by the appropriate conversion factor. These operations are probably as close as we can get within the least squares technique to the best fit by eye. If standard procedure is adopted for fitting by eye, scale units are chosen which cause the points to lie along a slope of approximately 1; in drawing the line the relative deviations of different points from possible lines are probably estimated by the shortest distance. Then the exact graphical slope is determined and reconverted to the original units.

Teissier (8) has pointed out that this line can also be derived by minimizing the sum of the areas of the right triangles whose hypotenuse is on the desired line and whose sides are the vertical and horizontal lines connecting the point with the desired line. The triangle used is

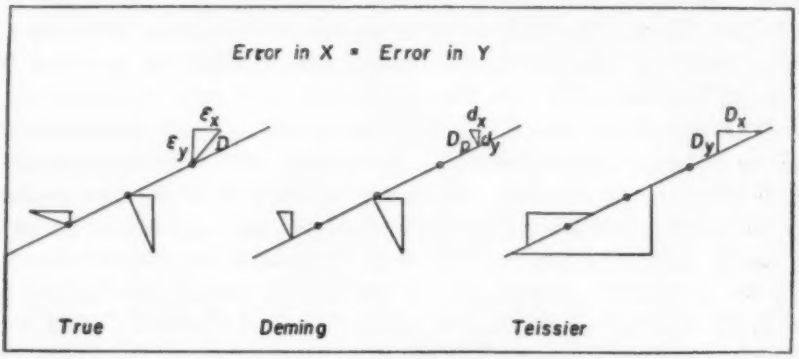


FIG. 6. FURTHER DEFINITIONS

indicated in Fig. 6. It does not represent an analysis of the error into x and y components such as is also shown in Fig. 6. Each side of the Teissier triangle is the maximal value of the error in this direction which would result if there were no error at all in the other direction.

This line has points in its favor: (1) It is an objective single-valued analytical equivalent for the best fit by eye. (2) In Figs. 5b, 5c and 5d there are three examples selected from a large set of interrelated variables and in Fig. 3c there is one example selected from another larger set of interrelated variables (rat bodily components and measurements). If such a set of variables is denoted by X', X'', X''', etc., one out of the set — say X' — may be selected as reference variable and least squares slopes k_y or k_a or $k_{p\text{-}\sigma}$ calculated for the relations $X'' - X'$, $X''' - X'$, $X'''' - X'$, etc. If from these slopes we calculate

the slopes of other relations not referred to X', such as the $X'' - X'''$ relation, the results will not be the same as the directly calculated least squares slopes k_y or k_x or $k_{y \cdot \sigma}$ for the $X'' - X'''$ relation, because the crossproduct terms $\Sigma x_1 x_2$, and $\Sigma x_1 x_3$ do not divide out to give $\Sigma x_2 x_3$. With the geometrical mean the slope is of the form $\sqrt{\Sigma y^2 / \Sigma x^2}$ with no crossproduct terms, so that within such a set of interrelated variables it is possible to get a set of solutions which is consistent with the relations known to exist among the various slopes. This is of considerable practical convenience. (Teissier has also discussed this point.)

This is not, however, a unique feature of this particular line; it will be true of many other possibilities — as for instance $|k| = \Sigma|y|/\Sigma|x|$ or $|k| = (\Sigma|y^3|/\Sigma|x^3|)^{1/3}$ or $k = (\Sigma y^4/\Sigma x^4)^{1/4}$. Also if one had a general least squares solution which was practicable and theoretically acceptable, such as Deming's, one could fit all these sets of data to their separate lines in one operation, imposing the additional conditions on the solutions that the required interrelations among slope values be met; one can continue to impose any other new conditions known to be required, in addition to the least squares condition, as long as there are not more conditions than data points. This is not offered as a practically useful solution, but simply to make it clear that the existence of exact interrelations among the true slopes of various sets of data provides no evidence with regard to whether the expression for the best estimate of a slope calculated for a single set of data should or should not have this property. If one doesn't know how to calculate the best estimate anyway, then it is frequently more convenient to use a solution with this property provided it meets the other known condition that it lie between the regression lines.

To the solution $k = (\Sigma y^2/\Sigma x^2)^{1/2}$ one can object that it cannot be the best solution. In general, when all the error is associated with Y, the line with slope $\Sigma xy/\Sigma x^2$ is the best; when all the error is associated with X, the line with slope $\Sigma y^2/\Sigma xy$ is the best; there is a variety of intermediate cases, not just one, and therefore the best solution requires a variety of intermediate lines, not just one. If it is known or suspected that the error is mostly in Y, but there is an appreciable, though small, contribution from X, the best solution must necessarily lie closer to $\Sigma xy/\Sigma x^2$ than it does to $\Sigma y^2/\Sigma xy$, and vice versa for the other case. The use of k_y, k_x or $\sqrt{k_y k_x}$ makes provision for only three gradations in the error distribution.

According to Teissier, this line with slope $(\Sigma y^2/\Sigma x^2)^{1/2}$ is satisfactory for all cases of representative samples of a population of X, Y points where X cannot be considered as dependent on Y, or vice versa. Then, he says, *"C'est vers cette ligne que tendraient les deux lignes de regression si les distributions marginales restant invariables, la corrélation tendait vers la perfection."* (8) The sense of this is not quite clear. If it means simply that the regression lines approach the line with slope $(\Sigma y^2/\Sigma x^2)^{1/2}$ as the correlation improves, it is true but is no evidence for this particular solution; these three lines and several others all approach the true line under these conditions. If it means that as the correlation is improved the numerical values of the regression slopes approach a stable (within sampling error) value calculatable for all degrees of correlation from the expression $(\Sigma y^2/\Sigma x^2)^{1/2}$ this would be the best reason for choosing this solution, but it is not true. As we have seen, if the error is practically entirely associated with Y (and this is a perfectly possible case of the situation with neither variable dependent on the other) it is the regression of Y on X which remains stable as the correlation is improved and both the regression of X on Y, and the line with slope $(\Sigma y^2/\Sigma x^2)^{1/2}$, approach this value. If the error is essentially in X, it is (in this same sense) towards the regression of X on Y that the two regression lines tend as the correlation is improved. If there is equal error with each variable, and Deming's solution is right, it is towards the line with slope k given by

$$k^2 + k\,\frac{(\Sigma x^2 - \Sigma y^2)}{\Sigma xy} - 1 = 0$$

that the two regression lines tend as the correlation is improved. If the line with slope $\sqrt{\Sigma y^2/\Sigma x^2}$ is actually approached by the two regression lines as the correlation is improved for some particular error distribution, this cannot be true for any other error distribution.

The method of averages. This procedure amounts to dividing the points into two groups, averaging X's and Y's within each group, and using the line connecting the two resulting mean points. As might be expected, and as Wald (17) has shown, the most favorable utilization of all the data results if the two mean points are equally reliable (i. e., equal or nearly equal numbers of points in each group) and as far away from each other as possible (i. e., the points are ordered, and all those on one side of the middle are in one group; those on the other side in the other group.) In general, the points will not be assigned to the same groups if they are ordered by Y as if they are ordered by X because of

error in Y or X or both. (If the points could be ordered by true X the result would be identical with ordering by true Y.) Having thus divided the observations into two groups, we now argue that the best line for each group should pass through the mean of the X's and Y's for that group (exactly as is true for the best line for a single group of points); and since the true line is the same for both sets of points, we may expect that the line through the two pairs of means will locate the true line within sampling error. On the same basis that the best line for the set corresponding to true points 1 through 9 should pass through the mean of all X's and Y's, the best line for the set from 1 through 4 should pass through the mean of the associated X's and Y's, and the best line for the set from 5 through 9 should pass through the mean of their associated X's and Y's. Since the true line for both sets is the same, these two means would locate the true line within sampling error.

Since we do not know how to order by true X or Y, Wald recommends as solution the line in which the assignment to groups is made on the basis of the observed X values; in terms of the variable $x = X - \overline{X}$, group 1 comprises all points with negatixe x, group 2 all points with x zero or positive. We will call this slope $k_{A\text{-}1}$. Wald does not discuss the other line, which may be designated as $k_{A\text{-}2}$, resulting from assignment to groups on the basis of the observed Y values, so that group 1 would comprise all points with negative y; group 2 all points with y zero or positive. These two lines are the same only when ordering with respect to X (or x) gives the same order of points as ordering with respect to Y (or y). When there is random scatter of points, $k_{A\text{-}1}$ approaches 0 and $k_{A\text{-}2}$ approaches ∞ , since ordering on x produces no ordering on y. The set of lines with slopes $k_{A\text{-}1}$ and $k_{A\text{-}2}$ behaves very similarly to the set of lines with slopes k_y and k_x. In several numerical examples it was found that $k_{A\text{-}1}$ (Wald's solution) $< k_y$ and $k_{A\text{-}2} > k_x$, indicating that $k_{A\text{-}1}$ and $k_{A\text{-}2}$ may approach the limiting values of 0 and ∞ more rapidly with deterioration in quality of fit than do k_y and k_x, but it is not clear whether this is generally true.

The slopes k_y and k_x will always differ as long as there are 3 or more points not all on one line. It is possible for $k_{A\text{-}1}$ and $k_{A\text{-}2}$ to be identical in value even when there is a fair number of points, but this identity merely means that ordering by observed x and by observed y is the same; it does not mean necessarily that this ordering is the same

as ordering by true x and y, and therefore that $k_{A\text{-}1} = k_{A\text{-}2}$ is an unbiased estimate of the slope.

Wald develops the criterion for validity of $k_{A\text{-}1}$ as follows: Let us take a reasonable maximal estimate of the possible error in x, called c. In the interval $x = +c$ to $x = -c$ will be all the points which owing to error will be differently ordered according to observed x than according to true x. If the number of points in this interval is sufficiently small, ordering by observed x is equivalent to ordering by true x. However, the only situation which accurately meets this condition is the situation with practically no error in x. If this is so, whether the error in y is large or small, $k_{A\text{-}1}$ should be an unbiased solution. Similarly, if there is practically no error in y, no matter what the error in x is, ordering by observed y will be practically the same as ordering by true y and then $k_{A\text{-}2}$ will be practically "independent of the errors," in Wald's phrase. The criterion of validity of $k_{A\text{-}1}$ and $k_{A\text{-}2}$ is slightly different than for k_y and k_x. In the former case, it hinges upon the requirement that the error in x (or in y) be small enough so that very few points will be in the doubtful interval. In the latter case, the requirement is that the error in x (or in y) be small enough in relation to the error in y (or in x) so that the deviations from the line can be taken in the y (or x) direction.

If an estimate of the error in x is not available from independent measurements, an upper limit can be placed upon it from the data themselves. The root mean square deviation of x from the regression of x on y is a maximal value for the root mean square error in x. This figure multiplied by three should be a safe value for c if the x errors are approximately normally distributed. If there are no x's in this interval, $k_{A\text{-}1}$ is an unbiased estimate of the slope. If there are x's in the interval, $k_{A\text{-}1}$ is a biased estimate which will in general, although not always, tend to be too low. But Wald points out a way to evaluate the full uncertainty in the value of the slope. If one considers all possible groupings of the points into two groups such that one group always contains the points on one side of the critical interval, and the other group always contains the points on the other side of the critical interval, but the points within the critcial interval are distributed in every possible fashion between the two groups, then a range of slope values can be calculated within which must lie the best estimate of the true slope. The confidence limits for the calculated slope value are represented by this range of values extended on both sides by the appropriate

function of the precision measure calculated from an unbiased estimate of the slope. The recommended solution $k_{A\text{-}1}$ with its confidence limits widened to include the full area of uncertainty due to inability to order the points independent of the error is quite analogous to the use of k_y with confidence limits widened to include the full area of uncertainty between k_y and k_x due to inability to measure the deviations independent of the error.

For cases with the error all in one variable, the appropriate least squares line should be more precise and the calculation provides more information; Wald's solution is easier to calculate; both solutions should be unbiased. When there is error in both variables and the error distribution is known, the appropriate least squares line provides the best solution because the only uncertainty in the slope is due to sampling error. When the error distribution is not known, there are two ways of estimating the area of uncertainty due to the error — the spread between k_y and k_x and the spread obtained by Wald's calculation. They will not in general be the same for a given set of data; we are justified in rejecting the larger estimate. However, the labor of calculating Wald's estimate rises so rapidly with increasing number of points in the critical interval that it is practical only for distributions with either few points, or very small error in one variable. When the error distribution is not known, there is apparently no way of calculating an unbiased estimate of the slope itself; any value within the area of uncertainty could be used, and it may then be convenient to use any suitable approximate solution which lies within this range and which may have other desirable properties.

SUMMARY

1. Assuming that errors are essentially random, and have a normal Gaussian probability distribution, it appears to follow that the best estimate of the true line must lie within (edges included) the area bounded by the two regression lines (least squares lines with deviations taken in the y direction, for one line, and in the x direction, for the other). Just where within this area the best solution lies depends upon the direction in which the points are regarded as deviating from the line, and this direction is difficult to fix in practice.

2. The frequent uncertainty in the choice of line for a particular problem makes the application of statistical significance tests to the slope or intercept values doubtful or misleading; unless the data points

are very few in number, the standard error of slope or intercept is much smaller than the real uncertainty in the value of slope or intercept, which is maximally in principle (and not infrequently in practice) the full difference between estimates from the two regression lines plus sampling error.

3. As the quality of fit to a linear trend is improved, either by reducing the error or by extending the range of the variables, the area bounded by the regression lines converges on the true line. The line of choice for any particular situation (both as estimate of the true line and for some types of prediction purposes) is the one for which successive values show no trend during such a converging process. In selected distributions, the way in which the various lines behave during such a converging process is roughly outlined.

4. The relation of the least squares lines to other approximate solutions is discussed, including the best fit by eye and the method of averages.

REFERENCES

1. Whittaker, E. T., and Robinson, G., The Calculus of Observations. Blackie and Son, *London,* 1924.
2. Mellor, J. W., Higher Mathematics for Students of Chemistry and Physics. Longmans Green, *London,* 1905, 1909.
3. Wright, W. T., and Hayford, J. F., The Adjustment of Observations. Van Nostrand, 1906, ed. 2.
4. Margenau, H., and Murphy, G. M., The Mathematics of Physics and Chemistry. Van Nostrand, 1943.
5. Birge, R. T., *Phys. Rev.,* 40, 207 (1932).
6. Ostwald-Luther, Physikochemische Messungen. Akadem. Verlag, *Leipzig,* 1925, ed. 4.
7. Ostwald-Luther, Physikochemische Messungen. 5th German ed., 1st American ed., Dover Publications, *New York,* 1943.
8. Teissier, G., "La Relation d'Allométrie, sa Signification Statistique et sa Logique", 1st Int. Biometrics Conf., Sept. 1947; to be published in *Biometrics,* March, 1948.
9. Uhler, H. S., *J. Opt. Soc. Amer.,* 7, 1043 (1923).
10. Camp, B. H., The Mathematical Part of Elementary Statistics. D. C. Heath, 1931, p. 152.
11. Yule, G. U., and Kendall, M. G., An Introduction to the Theory of Statistics. Charles Griffin and Co., *London,* 1937, ed. 11, footnote to page 314.
12. Greenwood, M., and Yule, G. U., *Proc. Roy. Soc. Medicine,* 8, 113 (1915).
13. Kenney, J. F., Mathematics of Statistics. Van Nostrand, 1941, pp. 141-142.

14. WIRTH, W., "Spezielle psychophysische Messmethoden," in E. Abderhalden's "Handbuch der biologischen Arbeitsmethoden." *Berlin*, 1920.
15. DEMING, W. E., Statistical Adjustment of Data. John Wiley, *New York*, 1943.
16. WORTHING, A. G., and GEFFNER, J., Treatment of Experimental Data. John Wiley, *New York*, 1943.
17. WALD, A., *Ann. Math. Stat.*, 11, 284 (1940).
18. BIRGE, R. T., *Rev. Mod. Phys.*, 1, 5 (1929).
19. MILLIKAN, R. A., *Phil. Mag.*, (6), 34, 1 (1917).
20. PHILLIPS, P., *Proc. Roy. Soc.*, A87, 48 (1912); JELLINEK, K., Lehrbuch der physikalischen chemie, vol. 1, Ferdinand Enke, *Stuttgart*, 1928, p. 497.
21. KINDLER, *Ann. d. Chem.*, 450, 1 (1926); 452, 90 (1927); 464, 278 (1928). HAMMETT, L. P., *J. Am. Chem. Soc.*, 59, 95 (1937).
22. HAMMETT, L. P., Physical Organic Chemistry. McGraw Hill, 1940. Reactions 3, 8 and 18 in Table II, p. 189.
23. BULL, H. B. and COOPER, J. A., in "Surface Chemistry." Amer. Assoc. Adv. Sci., *Washington*, 1943.
24. ZUCKER, T. F., and ZUCKER, L. M., *Am. J. Physiol.*, 146, 585 (1946).
25. WEYMOUTH, F. W., FIELD, J., and KLEIBER, M., *Proc. Soc. Exp. Biol. Med.*, 49, 367 (1942).
26. HAMMETT, L. P., *Trans. Far. Soc.*, 34, 156 (1938).

CORRELATION IN GROWTH

BY H. V. MUHSAM, Dr.-Ès-Sc.

INTRODUCTION

NORMS for the growth of school children have been prepared by different methods; but, as far as we know, all tables and charts compiled for this purpose are based on the computation of averages of corresponding measurements. Thus they represent series of averages rather than norms of development and it has never been proved that the normal child assumes in its development, one after the other, the means of heights and weights observed for a large sample of children.

Information on the normal development of school children can be obtained only from *seriatim* measurements, i. e., by observing the same individuals during their growth. It is difficult to collect data of this type but a large amount of information is contained in papers by Wilson (4), Palmer (2, 3), and others. The present study is an attempt to perform a complete analysis of the data included in Wilson's paper by correlation methods. Wilson calculated means and standard deviations of, and correlation coefficients between, the height or weight attained by a group of school girls at ages 7 to 16 years. It is difficult to interpret the physical meaning of the correlation coefficients given by Wilson, as the correlation between height or weight reached at different ages is due to two different factors: (a) the trivial fact that the height once attained by a child constitutes part of the height attained at any later age; and (b) the relationship between the height reached at a certain age and the increase during later years.

The same remark applies also to correlations in weight, although a decrease of weight may sometimes be observed, whereas no loss of height should occur.

METHOD USED

In a recent paper (1), a method has been worked out which permits us to separate the effect of the two factors mentioned under (a) and (b). According to this method, the correlation ($r_{i,\,i-j}$) between the

height attained at age (i) and the increase in height from age (i) to age (j) can be estimated from data of the type given in Wilson's paper by means of the following formula:

$$r_{i,\,i\text{-}j} = \frac{\sigma_j r_{ij} - \sigma_i}{(\sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j r_{ij})^{1/2}} \qquad (1)$$

where σ_i = standard deviation of heights at age (i),

σ_j = standard deviation of heights at age (j),

r_{ij} = coefficient of correlation between height at age (i) and height at age (j),

σ_i, σ_j and r_{ij} being known quantities. Similarly, the coefficient of correlation between the increase in height from age (i) to age (j) and that from age (k) to age (l) is given by the following formula:

$$r_{i\text{-}j,\,k\text{-}l} = \frac{\sigma_i\sigma_l r_{il} - \sigma_i\sigma_k r_{ik} - \sigma_j\sigma_k r_{jk} + \sigma_j\sigma_l r_{jl}}{(\sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j r_{ij})^{1/2}\,(\sigma_k^2 + \sigma_l^2 - 2\sigma_k\sigma_l r_{kl})^{1/2}} \qquad (2)$$

the symbols corresponding to those used in the previous formula.* The same formulae hold evidently also if weights replace heights.

The correlations included in Wilson's and in the present paper are computed by Pearson's product moment formula although it is doubtful whether all regressions are really linear. The high coefficients obtained by Wilson justify his method, but for the present study no direct proof of the linearity of regressions can be given, as the necessary information is not supplied by Wilson's publication. Some regressions of height and weight reached and the corresponding increase during

* [Formula (1) is easily derived by considering

$$x + y = z$$

where x, y, and z are random variables. From the relations

$$\sigma_x^2 + 2\sigma_x\sigma_y r_{xy} + \sigma_y^2 = \sigma_z^2,$$

$$\sigma_y^2 = \sigma_z^2 - 2\sigma_z\sigma_x r_{zx} + \sigma_x^2$$

we get at once

$$r_{xy} = \frac{\sigma_z r_{zx} - \sigma_x}{(\sigma_z^2 - 2\sigma_z\sigma_x r_{zx} + \sigma_x^2)^{1/2}}$$

which is the formula of the text. A similar procedure will give the author's second formula. Ed.]

the following year are shown by Palmer. In most cases, these regressions are not very far from linearity. In Table 1, correlation ratios

TABLE 1

Correlation coefficients (Wilson's data) and correlation ratios (Palmer's data) between height attained and increase during following year—girls

	YEARS OF AGE									
	6	7	8	9	10	11	12	13	14	15
Correlation coefficients (r)		.16	.27	.32	.36	.03	—.39	—.45	—.28	—.20
Correlation ratios (η)	.19	.11	.09	.14	.34	.28	.29	.49		

computed from Palmer's data are compared with correlation coefficients computed from Wilson's data. In some cases, the correlation ratio is considerably higher than the corresponding coefficient so that the regression might be rather curvilinear. Consequently, not too much importance should be attributed to numerical values of single correlation coefficients; yet the sign of a coefficient can in general be considered as fully reliable, and the general structure of the correlation patterns will also be affected to a small extent only, by the deviation from linearity of certain regressions.

RESULTS AND DISCUSSION

Table 2 shows coefficients of correlation between the height reached at each age from 6 to 15 years and the increase in height during the 1, 2, 3, etc., years (up to the age of 16) following the age attained; it shows also the corresponding correlations for weights. It will be noted from Table 2 that the correlations between height or weight attained at any age and the increase during one consecutive year is always relatively small; higher correlations are obtained when groups of years are used instead of individual years. The correlation between the height reached at ages from 7 to 9 years (or weight reached from 7 to 10 years) and the increase in height (or weight) during the following years is positive and increases with the number of "following years" included, if these do not exceed the age of 12. Thus, girls who, at

age 7, 8 or 9 years are relatively tall or heavy, grow, up to age 12, more than girls who are small or light at the younger age.

The correlation between height or weight attained at ages 7 to 9 (or even 10) years and the increase in height or weight during following years decreases, if the "following years" include ages over 12 years, and, for height, it becomes negative if ages over 14 years are

TABLE 2

Correlations between height (upper figures), weight (lower figures) and increase in height (or weight) from age x years to age y years

AGE y YEARS	AGE x YEARS								
	7	8	9	10	11	12	13	14	15
8	.16 .22								
9	.34 .40	.27 .16							
10	.36 .46	.35 .40	.32 .31						
11	.38 .53	.37 .53	.37 .50	.36 .47					
12	.34 .56	.36 .58	.30 .55	.23 .56	.03 .36				
13	.20 .52	.18 .55	.09 .49	—.07 .46	—.30 .27	—.39 .06			
14	—.01 .54	—.05 .51	—.16 .43	—.32 .30	—.49 .07	—.54 —.14	—.45 —.23		
15	—.13 .49	—.17 .43	—.27 .34	—.41 .17	—.55 —.06	—.58 —.23	—.49 —.27	—.28 —.13	
16	—.19 .47	—.22 .40	—.33 .30	—.45 .12	—.58 —.12	—.61 —.29	—.51 —.34	—.33 —.23	—.20 —.20

included. Thus, girls who are relatively tall at age 7 to 9 years grow, up to age 15 to 16 years, less than girls who are small at the age of 7 to 9 years.

This phenomenon of compensation in growth is more clearly exhibited in Table 3, which shows that, if the whole range of ages from

7 to 16 years is divided into two parts, the increase during the first part is in most cases negatively correlated with the increase during the second part. This negative correlation is strongest for the increase in height, if the whole range of ages is divided at age 11 or 12 years; whereas, for the increase in weight, the corresponding limit proves to be age 12 to 13 years. Thus girls who grow much during the years of age from 7 to 12, grow, during the years of age from 12 to 16,

TABLE 3

Coefficients of correlation between increase during groups of years, dividing range into two parts

	7–8 8–16	7–9 9–16	7–10 10–16	7–11 11–16	7–12 12–16	7–13 13–16	7–14 14–16	7–15 15–16
Height	—.13	—.26	—.52	—.68	—.68	—.45	—.11	+.10
Weight	+.12	+.11	—.03	—.23	—.38	—.40	—.25	—.22

less than girls who grow less during the previous period, and vice versa. It can therefore be said that growth during later childhood (7 to 12 years of age) is compensated, to a certain extent, by growth during puberty (12 to 16 years of age).

The figures of the right hand part of Table 2 also throw some light on the phenomenon of compensation: the height attained at age 12 or any later age is negatively correlated with the increase during following years, and the same is true for weights attained at age 13 and later. Girls who are relatively small at 12 or 13 years or later increase more, until the age of 16, than girls who are tall or heavy at the earlier age.

Additional and more detailed information on the phenomenon of compensation in growth can be obtained from Table 4, which shows, at the same time, that, from the point of view of growth, the range of ages from 7 years to 16 years falls naturally into two distinct periods, the period of later childhood, and the period of puberty. The increase in height during the year of age 7-8 is positively correlated with the increase during individual years of age from 9 to 12 years, but these coefficients are low. Higher coefficients are obtained between the increase during individual years of age between 8 and 11 years. Ages from 7 to 11 years, during which increases are highly intercor-

TABLE 4

Correlations between increases during single years; in height, upper figures; in weight, lower figures

	8–9	9–10	10–11	11–12	12–13	13–14	14–15	15–16
7–8	—.06 —.12	.25 .30	.15 .27	.10 .23	—.09 .10	—.12 —.09	—.14 —.14	—.08 —.05
8–9		.35 —.05	.24 .21	—.17 .09	—.30 —.05	—.26 —.10	—.24 —.07	—.26 —.08
9–10			.44 .21	—.10 .26	—.51 .04	—.48 —.28	—.40 —.23	—.33 —.11
10–11				.13 .17	—.44 .03	—.54 —.17	—.46 —.22	—.35 —.16
11–12					.13 .05	—.26 —.19	—.31 —.11	—.18 —.22
12–13						.53 .00	.34 .09	.21 —.09
13–14							.58 .18	.42 .04
14–15								.46 .01

related, thus constitute one period of growth; but it is not clear whether the year of age 7-8 belongs strictly to this period. Low negative correlations are observed between growth in height during the age 7-8 and that during single years of age from 12 to 16, whereas the increase during individual years of age from 8 to 12, and particularly the years from 9 to 11, is strongly (negatively) correlated with the increase during individual years of age from 12 to 16. The increase during each of the years of the latter period of growth is again positively correlated with the increase during any other year of the same period.

These conclusions are further confirmed by the factor analysis of the correlation coefficients of Table 4, which has been performed for the correlations in height only. Almost the whole variance can be accounted for by one factor (see Table 5). This factor shows high

saturations with increase during the ages from 8 to 11 years and again—negative ones—with increase during the ages from 12 to 16 years. But growth during the years of age 7-8 and 11-12 is not markedly affected by this factor.

The correlation pattern of annual increase in weight, also given in Table 4, is similar to that of increase in height, but less clear cut. For weights, as for heights, the increase during the year of age 7-8

TABLE 5

Factorial matrix of growth during single years of age

YEAR OF AGE	1ST FACTOR	2ND FACTOR
7- 8	.20	—.07
8- 9	.38	.30
9–10	.65	.28
10–11	.69	—.01
11–12	.12	—.57
12–13	—.59	—.32
13–14	—.82	.15
14–15	—.71	.30
15–16	—.53	—.15

is not closely correlated with the increase in period of late childhood (9 to 12 years of age); but with weights, the limit between the two periods of growth falls at age 12, or even 13 years, as compared with the age of 11 years observed in heights. The increases in weight during single years of age from 7 to 12 are negatively correlated with increases at the age from 13 to 16 years, which are positively intercorrelated. The year of age 12-13 cannot be attached definitely to the period of puberty growth, as far as increase in weight is concerned.

Most correlations between increases during single years of age are small, probably because of random fluctuations. Somewhat higher coefficients of correlation are shown in Table 6, which summarizes the information included in previous tables. The uppermost row of Table 6 shows correlations between the height and weight attained at age 7 years (or the total increase up to the age of 7 years) and the increase during triennial groups of age, and therefore summarizes clearly the features noticed in the first column of Table 2. The figures

TABLE 6

Correlation between height (upper figures) at age 7, weight (lower figures) and increase in height (or weight) during three-year periods; and correlations between increases during three-year periods

	7–10	10–13	13–16
7	.36 .46	—.07 .41	—.38 —.09
7–10		—.14 .39	—.50 —.39
10–13			—.17 —.06

given in the two lower rows of Table 6, confirm the results obtained from both the summary Table 3 and the detailed Table 4.

SUMMARY

The patterns of correlation in growth obtained by correlating weights and heights attained with increases during following years and by correlating increases, reveal three main facts:

(a) Growth in height and weight during different periods of the development is compensatory, i. e., girls growing more than average during a certain period grow less than average during another period, and vice versa.

(b) Growth of girls may be divided (for the range of ages investigated by Wilson) into two periods, one of late childhood, lasting from 7 or 8 years of age until an age between 11 and 13; the other, of puberty, lasts from 11 or 13 years of age until the end of the range of ages covered. Girls who grow much during one year of a period, grow much during the whole of that period, and those who grow little during one year, grow little during all years of the same period. During the period of late childhood, girls who are small or light at the beginning of the period grow less than girls who are tall or heavy at the beginning of the period, whereas for the period of puberty, the contrary is true.

(c) Although correlation patterns respective of growth in height

and in weight are similar in their general lines, a notable difference between them lies in the fact that, for growth in weight, the period of puberty growth and the effect of compensation begins one or two years later than for growth in height.

It is evident that the phenomenon of compensation cannot be accounted for in growth tables and charts based on averages, and that the lack of parallelism of growth in height and in weight causes additional difficulties in the establishment of norms of growth.

On the other hand, no final conclusions should be based on the correlations obtained by the indirect method used in the present study and from material which had not been collected for this particular purpose. It might be possible that some peculiarities of the correlation in growth patterns are due to certain characteristics of Wilson's material or the method of its collection. The standard deviations of height and weight distributions affect, for instance, the computed coefficients of correlation; but, as a matter of fact, the variance of these distributions is partly due to individual differences at the same age and partly to differences of age within the limits of one year. This second factor has different degrees of importance at various ages, as growth is much quicker at the age of 10 to 12 years than before and after this age. The hypothetical influence of this factor of variance could evidently not be accounted for in the present study.

Yet we do not believe that the general lines of the correlation in growth patterns obtained here must be ascribed only to shortcomings of methods and material used. In any event, there is no doubt that it would be worth while to collect and to analyse further data in order to confirm the results and conclusions of the present study and to obtain, at the same time, more detailed information on the norms of development of school children.

REFERENCES

1. Muhsam, H. V. (1947) Sur l'interpretation du coefficient de correlation. *Journ. Soc. Statist. Paris*, Vol. 88, No. 3-4.

2. Palmer, C. E. and Reed, L. J. (1935) Anthropometric studies of individual growth. I. Age, height and growth in height, elementary school children. Human Biology, Vol. 7, No. 3, pp. 319-334.

3. Palmer, C. E., Kawakami, R., and Reed, L. J. (1937) Anthropometric studies of individual growth. II. Age, weight and rate of growth in

weight, elementary school children. *Child Development.* Vol. 8, No. 1. pp. 47-61.

4. WILSON, E. B. (1935) Heights and weights of 275 public school girls for consecutive ages 7 to 16 years, inclusive. *Proc. Nat. Acad. Sc.,* Vol. 21, No. 12. pp. 633-634.

INDEX

Aboriginal populations, 83
Age distribution, Chinese population, 3
    Deaths in aboriginal and American Indian populations, 86
Age of family head and permanence of residence, 107
Age-specific fertility rates, for generations, 133
Allometry in monkeys, 53
Ateles, relative growth, 53

Birth rates, China, 3
    generation, 133
Blood pressure and chemical poisoning, 27
Bortkewitsch, L. von, 154

Cebus, relative growth, 53
Chemical exposure, statistical techniques in detecting, 27
Children, proportion in aboriginal and American Indian populations, 84
Chinese population, 1
Color, and permanence of residence, 91
Completed generation reproduction rates, 133
Cook, S. F., Survivorship in Aboriginal Populations, 83-89
Correlation, in growth, 260
    height and increase in height, 260
    height and weight of school children, 74
    weight and increase in weight, 260

"Das Gesetz der kleinen Zahlen," 154
Death rates, China, 3
Deaths, by age in aboriginal and American Indian populations, 85
Dextrality, sex differences, 21
Discordance as a measure of penetrance, 217

Eastern Health District, permanence of residence, 91
Education of family head, and permanence of residence, 110
Evaluation of slopes and intercepts of straight lines, 231

Family size, and permanence of residence, 112
Fertility, measured by completed generation reproduction rates, 133
Fitting straight lines, effects of errors in one or both variables, 231
Foulger chart, 42

Generation reproduction rates, 133
Group algebra, 186
Groups, social, 163
Growth, correlations in children, 260
  school children of two socio-economic groups, 68
  strength and weight, by sex, 12
Growth curves, monkeys, 53
  physical abilities, by sex, 12

Handedness, sex differences, 21
Height, and weight of Ottawa elementary school children of two socio-economic strata, 68
  correlated with increase in height, 260
  correlated with weight, school children, 74
Home ownership and permanence of residence, 100
HOPKINS, J. W., Height and Weight of Ottawa Elementary School Children of Two Socio-economic Strata, 68-82

Indians, deaths by age among American Indians, 85
Industrial preventive medicine, 27

JAFFE, A. J., Notes on the Rate of Growth of the Chinese Population, 1-11
JONES, HAROLD E., Sex Differences in Physical Abilities, 12-25

LASKER, GABRIEL W., Penetrance, Estimated by the Frequency of Unilateral Occurrences and by Discordance in Monozygotic Twins, (With an Editorial Note by Charles P. Winsor), 217-230
Laterality as a measure of penetrance, 217
Law of small numbers, 154
Least squares lines, 232
LIENAU, C. C., Quantitative Aspects of Organization, 163-216
LUMER, HYMAN, and SCHULTZ, ADOLPH H., Relative Growth of the Limb Segments and Tail in *Ateles geoffroyi* and *Cebus capucinus*, 53-67
LUYKX, H. M. C., Family Studies in the Eastern Health District. IV. Permanence of Residence with Respect to Various Family Characteristics, 91-132

Menarche related to growth in strength, 19
Migration of families, 91
Monkeys, relative growth of limb segments and tail, 53
MUHSAM, H. V., Correlation in Growth, 260-269

Nativity of family head, and permanence of residence, 109
Negative binomial ("Unequal") distribution, 49
Negro, permanence of residence, 91

Organization, measure of, 163

Peck order, 202
Penetrance, estimated by the frequency of unilateral occurrences and by discordance in monozygotic twins, 217
Permanence of residence with respect to various family characteristics, 91
Poisson frequency distribution, 154
Population, aboriginal survivorship, 83
China, 1
Preventive medicine, industrial, 27
Probability theory in contest, 190

Quantitative aspects of organization, 163

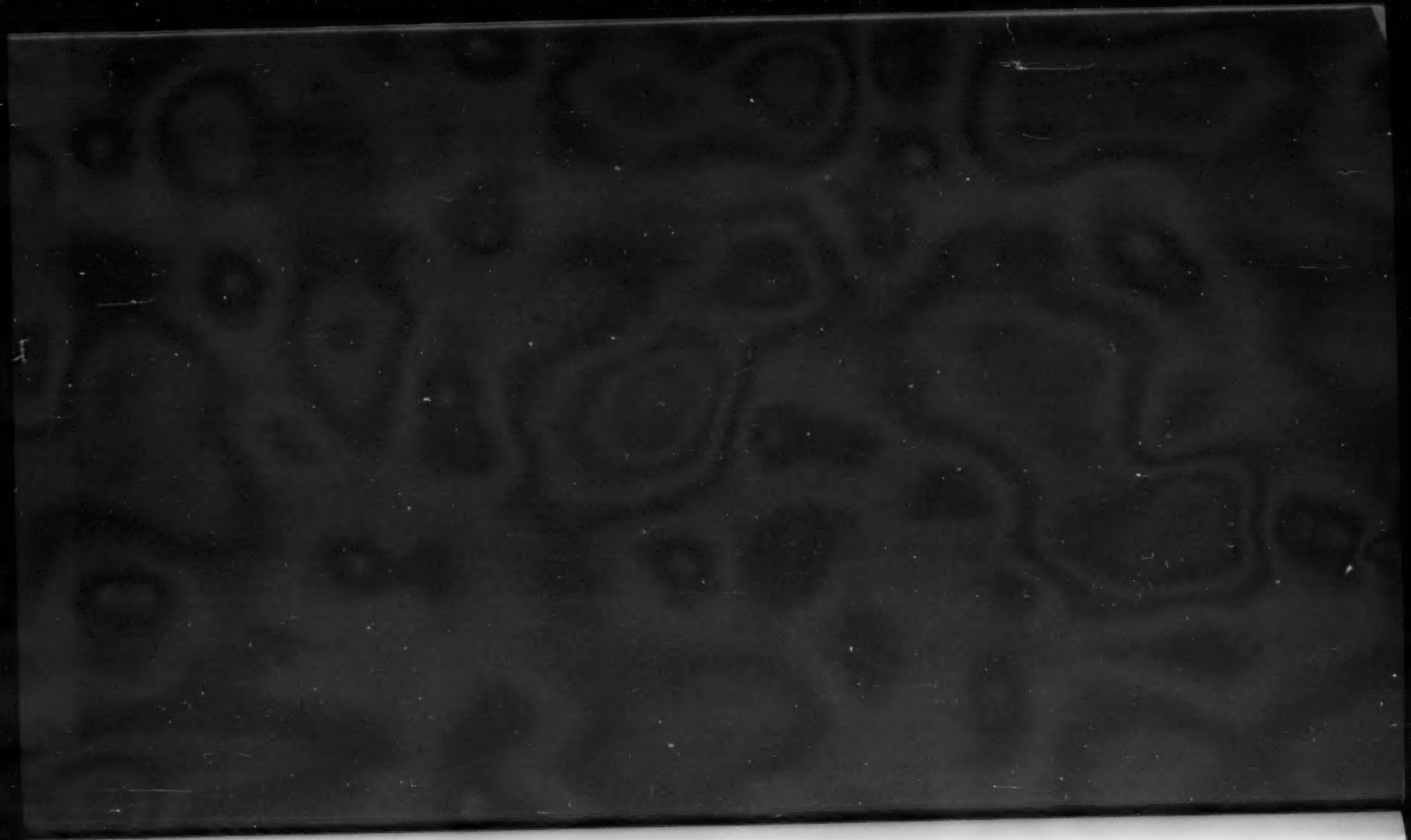
Racing, horse, 181
Rate of growth of the Chinese population, 1
Regression lines, methods of fitting, 231
Relative growth of the limb segments and tail in *Ateles geoffroyi* and *Cebus capucinus*, 53
Relief status, and permanence of residence, 117
Reproduction rates, China, 7
generation, 133
Right handedness, sex differences, 21

Sampling problems in measuring penetrance, 223
Schultz, Adolph H., 53
Sex differences in physical abilities, 12
Shewhart control chart technique, 31
Size of family and permanence of residence, 112
Socio-economic status, and growth, 68
and permanence of residence, 105
Statistical approaches to some problems of industrial preventive medicine, 27
Strength, sex differences in growth, 12
Supreme Court, agreement of Justices of, 212
Survival rates, generation, 133
Survivorship in aboriginal populations, 83

Twins, discordance in monozygotic, 217

Wage earners in family, and permanence of residence, 114
Weight, and increase in weight, correlations, 260
school children of two socio-economic groups, 68
sex differences in growth, 12
WINSOR, CHARLES P., Quotations: "Das Gesetz der kleinen Zahlen," 154-161
Winsor, Charles P., 217
WOOFTER, T. J., Completed Generation Reproduction Rates, 133-153

ZAPP, JOHN A., JR., Statistical Approaches to Some Problems of Industrial Preventive Medicine, 27-52
ZUCKER, LOIS M., Evaluation of Slopes and Intercepts of Straight Lines, 231-259



Human Biology

a record of research

VOLUME 19 DECEMBER, 1947 NUMBER 4

CONTENTS

QUANTITATIVE ASPECTS OF ORGANIZATION. *C. C. Lieaos* 163

PENETRANCE, ESTIMATED BY THE FREQUENCY OF UNILATERAL OCCURRENCES AND BY DISCORDANCE IN MONOZYGOTIC TWINS. *Gabriel W. Lasker* With an *Editorial Note by Charles B. Wilson* 237

EVALUATION OF SLOPES AND INTERCEPTS OF STRAIGHT LINES. *Lois M. Zucker* 231

CORRELATION IN GROWTH. *H. V. Muhsam* 260

INDEX 271

THE JOHNS HOPKINS PRESS, *Publishers*
BALTIMORE—18, MARYLAND, U. S. A.